BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, 15/62, C07K 14/20, A61K 39/02, G01N 33/50, C07K 16/12

A1

(11) International Publication Number: **WO 95/12676**

(43) International Publication Date: 11 May 1995 (11.05.95)

(21) International Application Number: PCT/US94/12352

(22) International Filing Date: 27 October 1994 (27.10.94)

(30) Priority Data:

08/148,191	1 November 1993 (01.11.93)	US
08/235,836	29 April 1994 (29.04.94)	US

(71) Applicant: ASSOCIATED UNIVERSITIES, INC. [US/US]; Suite 730, 1400 16th Street N.W., Washington, DC 20036 (US).

(72) Inventors: DUNN, John, J.; 5 Mott Drive, Bellport, NY 11713 (US). LUFT, Benjamin, J.; 223 Lincoln Avenue, Port Jefferson, NY 11777 (US).

(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).

(81) Designated States: AU, CA, FI, JP, NO, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: CHIMERIC PROTEINS COMPRISING BORRELIA POLYPEPTIDES: USES THEREFOR

(57) Abstract

Novel chimeric nucleic acids, encoding chimeric *Borrelia* proteins consisting of at least two antigenic polypeptides from corresponding and/or non-corresponding proteins from the same and/or different species of *Borrelia*, are disclosed. Chimeric proteins encoded by the nucleic acid sequences are also disclosed. The chimeric proteins are useful as vaccine immunogens against Lyme borreliosis, as well as for immunodiagnostic reagents.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

CHIMERIC PROTEINS COMPRISING BORRELIA POLYPEPTIDES; USES THEREFOR

Background of the Invention

Lyme borreliosis is the most common tick-borne infectious disease in North America, Europe, and northern Asia. The causative bacterial agent of this disease, *Borrelia burgdorferi*, was first isolated and cultivated in 1982 (Burgdorferi, W.A. *et al.*, Science 216: 1317-1319 (1982); Steere, A.R. *et al.*, N. Engl. J. Med. 308: 733-740 (1983)). With that discovery, a wide array of clinical syndromes, described in both the European and American literature since the early 20th century, could be attributed to infection by *B. burgdorferi* (Afzelius, A., Acta Derm. Venereol. 2: 120-125 (1921); Bannwarth, A., Arch. Psychiatr. Nervenkrankh. 117: 161-185 (1944); Garin, C. and A. Bujadouz, J. Med. Lyon 71: 765-767 (1922); Herxheimer, K. and K. Hartmann, Arch. Dermatol. Syphilol. 61: 57-76, 255-300 (1902)).

The immune response to *B. burgdorferi* is characterized by an early, prominent, and persistent humoral response to the end of lagellar protein, p41 (fla), and to a protein constituent of the protoplasmic cylinder, p93 (Szczepanski, A., and J.L. Benach, Microbiol. Rev. 55:21 (1991)). The p41 flagellin antigen is an immunodominant protein; however, it shares significant homology with flagellins of other microorganisms and therefore is highly cross reactive. The p93 antigen is the largest immunodominant antigen of *B. burgdorferi*. Both the p41 and p93 proteins are physically cryptic antigens, sheathed from the immune system by an outer membrane whose major protein constituents are the outer surface proteins A and B

(OspA and OspB). OspA is a basic lipoprotein of approximately 31 kd, which is encoded on a large linear plasmid along with OspB, a basic lipoprotein of approximately 34 kd (Szczepanski, A., and J.L. Benach, Microbiol. Rev. 55:21 (1991)). Analysis of isolates of *B. burgdorferi* obtained from North America and Europe has demonstrated that OspA has antigenic variability, and that several distinct groups can be serologically and genotypically defined (Wilske, B., *et al.*, World J. Microbiol. 7: 130 (1991)). Other *Borrelia* proteins demonstrate similar antigenic variability. Surprisingly, the immune response to these outer surface proteins tends to occur late in the disease, if at all (Craft, J. E. *et al.*, J. Clin Invest. 78: 934-939 (1986); Dattwyler, R.J. and B.J. Luft, Rheum. Clin. North Am. 15: 727-734 (1989)). Furthermore, patients acutely and chronically infected with *B. burgdorferi* respond variably to the different antigens, including OspA, OspB, OspC, OspD, p39, p41 and p93.

Vaccines against Lyme borreliosis have been attempted. Mice immunized with a recombinant form of OspA are protected from challenge with the same strain of *B. burgdorferi* from which the protein was obtained (Fikrig, E., *et al.*, Science 250: 553-556 (1990)). Furthermore, passively transferred anti-OspA monoclonal antibodies (Mabs) have been shown to be protective in mice, and vaccination with a recombinant protein induced protective immunity against subsequent infection with the homologous strain of *B.burgdorferi* (Simon, M.M., *et al.*, J. Infect. Dis. 164: 123 (1991)). Unfortunately, immunization with a protein from one strain does not necessarily confer resistance to a heterologous strain (Fikrig, E. *et al.*, J. Immunol. 7: 2256-1160 (1992)), but rather, is limited to the homologous 'species' from which the protein was prepared. Furthermore,

immunization with a single protein from a particular strain of *Borrelia* will not confer resistance to that strain in all individuals. There is considerable variation displayed in OspA and OspB, as well as p93, including the regions conferring antigenicity. Therefore, the degree and frequency of protection from vaccination with a protein from a single strain depend upon the response of the immune system to the particular variation, as well as the frequency of genetic variation in *B. burgdorferi*. Currently, a need exists for a vaccine which provides immunogenicity across species and to more epitopes within a species, as well as immunogenicity against more than one protein.

Summary of the Invention

The current invention pertains to chimeric *Borrelia* proteins which include two or more antigenic *Borrelia* polypeptides which do not occur naturally (in nature) in the same protein in *Borrelia*, as well as the nucleic acids encoding such chimeric proteins. The antigenic polypeptides incorporated in the chimeric proteins are derived from any *Borrelia* protein from any strain of *Borrelia*, and include outer surface protein (Osp) A, OspB, OspC, OspD, p12, p39, p41, p66, and p93. The proteins from which the antigenic polypeptides are derived can be from the same strain of *Borrelia*, from different strains, or from combinations of proteins from the same and from different strains. If the proteins from which the antigenic polypeptides are derived are OspA or OspB, the antigenic polypeptides can be derived from either the portion of the OspA or OspB protein present between the amino terminus and the conserved tryptophan of the protein (referred to as a proximal portion), or the portion of the OspA or OspB protein present between the conserved tryptophan of the protein

and the carboxy terminus (referred to as a distal portion). Particular chimeric proteins, and the nucleotide sequences encoding them, are set forth in Figures 23-37 and 43-46.

The chimeric proteins of the current invention provide antigenic polypeptides of a variety of *Borrelia* strains and/or proteins within a single protein. Such proteins are particularly useful in immunodiagostic assays to detect the presence of antibodies to native *Borrelia* in potentially infected individuals as well as to measure T-cell reactivity, and can therefore be used as immunodiagnostic reagents. The chimeric proteins of the current invention are additionally useful as vaccine immunogens against *Borrelia* infection.

For a better understanding of the present invention together with other and further objects, reference is made to the following description, taken together with the accompanying drawings.

Brief Description of the Drawings

Figure 1 summarizes peptides and antigenic domains localized by proteolytic and chemical fragmentation of OspA.

Figure 2 is a comparison of the antigenic domains depicted in Figure 1, for OspA in nine strains of *B. burgdorferi*.

Figure 3 is a graph depicting a plot of weighted polymorphism versus amino acid position among 14 OspA variants. The marked peaks are: a) amino acids 132-145; b) amino acids 163-177; c) amino acids 208-221. The lower dotted line at polymorphism value 1.395 demarcates statistically significant excesses of polymorphism at p = 0.05. The upper dotted line at 1.520 is the same, except that the first 29 amino acids at the monomorphic N-terminus have been removed from the original analysis.

Figure 4 depicts the amino acid alignment of residues 200 through 220 for OspAs from strains B31 and K48 as well as for the site-directed mutants 613, 625, 640, 613/625, and 613/640. Arrow indicates Trp216. Amino acid changes are underlined.

Figure 5 is a helical wheel projection of residues 204-217 of B31 OspA. Capital letters indicate hydrophobic residues; lower case letters indicate hydrophilic residues; +/- indicate positively/negatively charged residues. Dashed line indicates division of the alpha-helix into hydrophobic arc (above the line) and polar arc (below the line). Adapted from France *et al.* (Biochem. Biophys. Acta 1120: 59 (1992)).

Figure 6 depicts a phylogenic tree for strains of *Borrelia* described in Table I. The strains are as follows: 1 = B31; 2 = Pka1; 3 = ZS7; 4 = N40; 5 = 25015; 6 = K48; 7 = DK29; 8 = PHei; 9 = Ip90; 10 = PTrob; 11 = ACAI; 12 = PGau; 13 = Ip3; 14 = PBo; 15 = PKo.

Figure 7 depicts the nucleic acid sequence of OspA-B31 (SEW ID NO. 6), and the encoded protein sequence (SEQ ID NO. 7).

Figure 8 depicts the nucleic acid sequence of OspA-K48 (SEQ ID NO. 8), and the encoded protein sequence (SEQ ID NO. 9).

Figure 9 depicts the nucleic acid sequence of OspA-PGau (SEQ ID NO. 10), and the encoded protein sequence (SEQ ID NO. 11).

Figure 10 depicts the nucleic acid sequence of OspA-25015 (SEQ ID NO. 12), and the encoded protein sequence (SEQ ID NO. 13).

Figure 11 depicts the nucleic acid sequence of OspB-B31 (SEQ ID NO. 21), and the encoded protein sequence (SEQ ID NO. 22).

Figure 12 depicts the nucleic acid sequence of OspC-B31 (SEQ ID NO. 29), and the encoded protein sequence (SEQ ID NO. 30).

Figure 13 depicts the nucleic acid sequence of OspC-K48 (SEQ ID NO. 31), and the encoded protein sequence (SEQ ID NO. 32).

Figure 14 depicts the nucleic acid sequence of OspC-PKo (SEQ ID NO. 33), and the encoded protein sequence (SEQ ID NO. 34).

Figure 15 depicts the nucleic acid sequence of OspC-pTrob (SEQ ID NO. 35) and the encoded protein sequence (SEQ ID NO. 36).

Figure 16 depicts the nucleic acid sequence of p93-B31 (SEQ ID NO. 65) and the encoded protein sequence (SEQ ID NO. 66).

Figure 17 depicts the nucleic acid sequence of p93-K48 (SEQ ID NO. 67).

Figure 18 depicts the nucleic acid sequence of p93-PBo (SEQ ID NO. 69).

Figure 19 depicts the nucleic acid sequence of p93-pTrob (SEQ ID NO. 71).

Figure 20 depicts the nucleic acid sequence of p93-pGau (SEQ ID NO. 73).

Figure 21 depicts the nucleic acid sequence of p93-25015 (SEQ ID NO. 75).

Figure 22 depicts the nucleic acid sequence of p93-pKo (SEQ ID NO. 77).

Figure 23 depicts the nucleic acid sequence of the OspA-K48/OspA-PGau chimer (SEQ ID NO. 85) and the encoded chimeric protein sequence (SEQ ID NO. 86).

Figure 24 depicts the nucleic acid sequence of the OspA-B31/OspA-PGau chimer (SEQ ID NO. 88) and the encoded chimeric protein sequence (SEQ ID NO. 89).

Figure 25 depicts the nucleic acid sequence of the OspA-B31/OspA-K48 chimer (SEQ ID NO. 91) and the encoded chimeric protein sequence (SEQ ID NO. 92).

Figure 26 depicts the nucleic acid sequence of the OspA-B31/OspA-25015 chimer (SEQ ID NO. 94) and the encoded chimeric protein sequence (SEQ ID NO. 95).

Figure 27 depicts the nucleic acid sequence of the OspA-K48/OspA-B31/OspA-K48 chimer (SEQ ID NO. 97) and the encoded chimeric protein sequence (SEQ ID NO. 98).

Figure 28 depicts the nucleic acid sequence of the OspA-B31/OspA-K48/OspA-B31/OspA-K48 chimer (SEQ ID NO. 100) and the encoded chimeric protein sequence (SEQ ID NO. 101).

Figure 29 depicts the nucleic acid sequence of the OspA-B31/OspB-B31 chimer (SEQ ID NO. 103) and the encoded chimeric protein sequence (SEQ ID NO. 104).

Figure 30 depicts the nucleic acid sequence of the OspA-B31/OspB-B31/OspC-B31 chimer (SEQ ID NO. 106) and the encoded chimeric protein sequence (SEQ ID NO. 107).

Figure 31 depicts the nucleic acid sequence of the OspC-B31/OspA-B31/OspB-B31 chimer (SEQ ID NO. 109) and the encoded chimeric protein sequence (SEQ ID NO. 110).

Figure 32 depicts the nucleic acid sequence of the OspA-B31/p93-B31 chimer (SEQ ID NO. 111) and the encoded chimeric protein sequence (SEQ ID NO. 112).

Figure 33 depicts the nucleic acid sequence of the OspB-B31/p41-B31 (122-234) chimer (SEQ ID NO. 113) and the encoded chimeric protein sequence (SEQ ID NO. 114).

Figure 34 depicts the nucleic acid sequence of the OspB-B31/p41-B31 (122-295) chimer (SEQ ID NO. 115) and the encoded chimeric protein sequence (SEQ ID NO. 116).

Figure 35 depicts the nucleic acid sequence of the OspB-B31/p41-B31 (140-234) chimer (SEQ ID NO. 117) and the encoded chimeric protein sequence (SEQ ID NO. 118).

Figure 36 depicts the nucleic acid sequence of the OspB-B31/p41-B31 (140-295) chimer (SEQ ID NO. 119) and the encoded chimeric protein sequence (SEQ ID NO. 120).

Figure 37 depicts the nucleic acid sequence of the OspB-B31/p41-B31 (122-234)/OspC-B31 chimer (SEQ ID NO. 121) and the encoded chimeric protein sequence (SEQ ID NO. 122).

Figure 38 depicts an alignment of the nucleic acid sequences for OspC-B31 (SEQ ID NO. 29), OspC-PKo (SEQ ID NO. 33), OspC-pTrob (SEQ ID NO. 35), and OspC-K48 (SEQ ID NO. 31). Nucleic acids which are identical to those in the lead nucleic acid sequence (here, OspC-B31) are represented by a period (.); differing nucleic acids are shown in lower case letters.

Figure 39 depicts an alignment of the nucleic acid sequences for OspD-pBO (SEQ ID NO. 123), OspD-PGau (SEq ID NO. 124), OspD-DK29 (SEQ ID NO. 125), and OspD-K48 (SEQ ID NO. 126). Nucleic acids which are identical to those in the lead nucleic acid sequence (here, OspD-pBo) are represented by a period (.); differing nucleic acids are shown in lower case letters.

Figure 40 depicts the nucleic acid sequence of p41-B31 (SEq ID NO. 127) and then encoded protein sequence (SEQ ID NO. 128).

Figure 41 depicts an alignment of the nucleic acid sequences for p41-B31 (SEQ ID NO. 127), p41-pKa1 (SEQ ID NO. 129), p41-PGau (SEQ ID NO. 51), p41-PBo (SEQ ID NO. 130), p41-DK29 (SEQ ID NO. 53), and p41-PKo (SEQ ID NO. 131). Nucleic acids which are identical to those in the lead nucleic acid sequence (here, p41-B31) are represented by a period (.); differing nucleic acids are shown in lower case letters.

Figure 42 depicts an alignment of the nucleic acid sequences for OspA-B31 (SEQ ID NO. 6), OspA-pKa1 (SEQ ID NO. 132), OspA-N40 (SEQ ID NO. 133), OspA-ZS7 (SEQ ID

NO. 134), OspA-25015 (SEQ ID NO. 12), OspA-pTrob (SEQ ID NO. 135), OspA-K48 (SEQ ID NO. 8), OspA-Hei (SEQ ID NO. 136), OspA-DK29 (SEQ ID NO. 49), OSpA-Ip90 (SEQ ID NO. 50), OspA-pBo (Seq ID NO. 55), OspA-Ip3 (SEQ ID NO. 56), OspA-PKo (SEQ ID NO. 57), OspA-ACAI (SEQ ID NO. 58), and OspA-PGau (SEQ ID NO. 10). Nucleic acids which are identical to those in the lead nucleic acid sequence (here, OspA-B31) are represented by a period (.); differing nucleic acids are shown in lower case letters.

Figure 43 depicts the nucleic acid sequence of the OspA-Tro/OspA-Bo chimer (SEQ ID NO. 137) and the encoded chimeric protein sequence (SEQ ID NO. 138).

Figure 44 depicts the nucleic acid sequence of the OspA-PGau/OspA-Bo chimer (SEQ ID NO. 139) and the encoded chimeric protein sequence (SEQ ID NO. 140).

Figure 45 depicts the nucleic acid sequence of the OspA-B31/OspA-PGau/OspA-B31/OspA-K48 chimer (SEQ ID NO. 141) and the encoded chimeric protein sequence (SEQ ID NO. 142).

Figure 46 depicts the nucleic acid sequence of the OspA-PGau/OspA-B31/OspA-K48 chimer (SEQ ID NO. 143) and the encoded chimeric protein sequence (SEQ ID NO. 144).

Detailed Description of the Invention

The current invention pertains to chimeric proteins comprising antigenic *Borrelia* polypeptides which do not occur in nature in the same *Borrelia* protein. The chimeric proteins are a combination of two or more antigenic polypeptides derived from *Borrelia* proteins. The antigenic polypeptides can be derived from different proteins from the same species of *Borrelia*, or different proteins from different *Borrelia* species, as well as from corresponding proteins from different species. As used herein, the term "chimeric protein" describes a protein comprising two or more polypeptides which are

derived from corresponding and/or non-corresponding native *Borrelia* protein. A polypeptide "derived from" a native *Borrelia* protein is a polypeptide which has an amino acid sequence the same as an amino acid sequence present in a *Borrelia* protein, an amino acid sequence equivalent to the amino acid sequence of a naturally occurring *Borrelia* protein, or an amino acid sequence substantially similar to the amino acid sequence of a naturally occurring *Borrelia* protein (e.g., differing by few amino acids) such as when a nucleic acid encoding a protein is subjected to site-directed mutagenesis. "Corresponding" proteins are equivalent proteins from different species or strains of *Borrelia*, such as outer surface protein A (OspA) from strain B31 and OspA from strain K48. The invention additionally pertains to nucleic acids encoding these chimeric proteins.

As described below, Applicants have identified two separate antigenic domains of OspA and OspB which flank the sole conserved tryptophan present in OspA and in OspB. These domains share cross-reactivity with different genospecies of *Borrelia*. The precise amino acids responsible for antigenic variability were determined through site-directed mutagenesis, so that proteins with specific amino acid substitutions are available for the development of chimeric proteins. Furthermore, Applicants have identified immunologically important hypervariable domains in OspA proteins, as described below in Example 2. The first hypervariable domain of interest for chimeric proteins, Domain A, includes amino acid residues 120-140 of OspA, the second hypervariable domain, Domain B, includes residues 150-180 and the third hypervariable domain, Domain C, includes residues 200-216 or 217 (depending on the position of the sole conserved tryptophan residue in the OspA of that particular species of *Borrelia*) (see Figure

3). In addition, Applicants have sequenced the genes for several *Borrelia* proteins.

These discoveries have aided in the development of novel recombinant *Borrelia* proteins which include two or more amino acid regions or sequences which do not occur in the same *Borrelia* protein in nature. The recombinant proteins comprise polypeptides from a variety of *Borrelia* proteins, including, but not limited to, OspA, OspB, OspC, OspD, p12, p39, p41, p66, and p93. Antigenically relevant polypeptides from each of a number of proteins are combined into a single chimeric protein.

In one embodiment of the current invention, chimers are now available which include antigenic polypeptides flanking a tryptophan residue. The antigenic polypeptides are derived from either the proximal portion from the tryptophan (the portion of the OspA or OspB protein present between the amino terminus and the conserved tryptophan of the protein), or the distal portion from the tryptophan (the portion of the OspA or OspB protein present between the conserved tryptophan of the protein and the carboxy terminus) in OspA and/or OspB. The resultant chimers can be OspA-OspA chimers (i.e., chimers incorporating polypeptides derived from OspA from different strains of *Borrelia*), OspA-OspB chimers, or OspB-OspB chimers, and are constructed such that amino acid residues amino-proximal to an invariant tryptophan are from one protein and residues carboxy-proximal to the invariant tryptophan are from the other protein. For example, one available chimer consists of a polypeptide derived from the amino-proximal region of OspA from strain B31, followed by the tryptophan residue, followed by a polypeptide derived from the carboxy-proximal region of OspA from strain K48 (SEQ ID NO. 92). Another available chimer includes a

polypeptide derived from the amino-proximal region of OspA from strain B31, and a polypeptide derived from the carboxy-proximal region of OspB from strain B31 (SEQ ID NO. 104). If the polypeptide proximal to the tryptophan of these chimeric proteins is derived from OspA, the proximal polypeptide can be further subdivided into the three hypervariable domains (Domains A, B, and C), each of which can be derived from OspA from a different strain of *Borrelia*. These chimeric proteins can further comprise antigenic polypeptides from another protein, in addition to the antigenic polypeptides flanking the tryptophan residue.

In another embodiment of the current invention, chimeric proteins are available which incorporate antigenic domains of two or more *Borrelia* proteins, such as Osp proteins (Osp A, B, C and/or D) as well as p12, p39, p41, p66, and/or p93.

The chimers described herein can be produced so that they are highly soluble, hyper-produced in *E. coli*, and non-lipidated. In addition, the chimeric proteins can be designed to end in an affinity tag (His-tag) to facilitate purification. The recombinant proteins described herein have been constructed to maintain high levels of antigenicity. In addition, recombinant proteins specific for the various genospecies of *Borrelia* that cause Lyme disease are now available, because the genes from each of the major genospecies have been sequenced; the sequences are set forth below. These recombinant proteins with their novel biophysical and antigenic properties will be important diagnostic reagent and vaccine candidates.

The chimeric proteins of the current invention are advantageous in that they retain specific reactivity to monoclonal and polyclonal antibodies against wild-type *Borrelia* proteins, are immunogenic, and inhibit the

growth or induce lysis of *Borrelia in vitro*. Furthermore, in some embodiments, the proteins provide antigenic domains of two or more *Borrelia* strains and/or proteins within a single protein. Such proteins are particularly useful in immuno-diagostic assays. For example, proteins of the present invention can be used as reagents in assays to detect the presence of antibodies to native *Borrelia* in potentially infected individuals. These proteins can also be used as immunodiagnostic reagents, such as in dot blots, Western blots, enzyme linked immunosorbed assays, or agglutination assays. The chimeric proteins of the present invention can be produced by known techniques, such as by recombinant methodology, polymerase chain reaction, or mutagenesis.

Furthermore, the proteins of the current invention are useful as vaccine immunogens against *Borrelia* infection. Because *Borrelia* has been shown to be clonal, a protein comprising antigenic polypeptides from a variety of *Borrelia* proteins and/or species, will provide immunoprotection for a considerable time when used in a vaccine. The lack of significant intragenic recombination, a process which might rapidly generate novel epitopes with changed antigenic properties, ensures that *Borrelia* can only change antigenic type by accumulating mutational change, which is slow when compared with recombination in generating different antigenic types. The chimeric protein can be combined with a physiologically acceptable carrier and administered to a vertebrate animal through standard methods (e.g., intravenously or intramuscularly, for example).

The current invention is illustrated by the following Examples, which are not to be construed to be limiting in any way.

Example 1. Purification of *Borrelia burgorferi* Outer Surface Protein A and Analysis of Antibody Binding Domains

This example details a method for the purification of large amounts of native outer surface protein A (OspA) to homogeneity, and describes mapping of the antigenic specificities of several anti-OspA MAbs. OspA was purified to homogeneity by exploiting its resistance to trypsin digestion. Intrinsic labeling with ^{14}C-palmitic acid confirmed that OspA was lipidated, and partial digestion established lipidation at the amino-terminal cysteine of the molecule.

The reactivity of seven anti-OspA murine monoclonal antibodies to nine different *Borrelia* isolates was ascertained by Western blot analysis. Purified OspA was fragmented by enzymatic or chemical cleavage, and the monoclonal antibodies were able to define four distinct immunogenic domains (see Figure 1). Domain 3, which included residues 190-220 of OspA, was reactive with protective antibodies known to agglutinate the organism *in vitro*, and included distinct specificities, some of which were not restricted to a genotype of *B. burgdorferi*.

A. Purification of Native OspA

Detergent solubilization of *B. burgdorferi* strips the outer surface proteins and yields partially-purified preparations containing both OspA and outer surface protein B (Osp B) (Barbour, A.G. *et al.*, Infect. Immun. 52 (5): 549-554 (1986); Coleman, J.L. and J.L. Benach, J Infect. Dis. 155 (4): 756-765 (1987); Cunningham, T.M. *et al.*, Ann. NY Acad. Sci. 539: 376-378 (1988); Brandt, M.E. *et al.*, Infect. Immun. 58: 983-991 (1990); Sambri, V. and R. Cevenini, Microbiol. 14:307-314 (1991)). Although both OspA and OspB are sensitive to proteinase K digestion, in contrast to OspB, OspA is resistant to cleavage by trypsin (Dunn, J. *et al.*, Prot. Exp. Purif. 1: 159-168 (1990); Barbour, A.G. *et al.*, Infect. Immun. 45:94-100 (1984)). The relative insensitivity to trypsin is surprising in view of the fact that Osp A has a high (16% for B31) lysine content, and may relate to the relative configuration of Osp A and B in the outer membrane.

Intrinsic Radiolabeling of Borrelia

Labeling for lipoproteins was performed as described by Brandt *et al.* (Infect. Immun. 58:983-991 (1990)). ^{14}C-palmitic acid (ICN, Irvine, California) was added to the BSK II media to a final concentration of 0.5 μCi per milliliter (ml). Organisms were cultured at 34°C in this medium until a density of 10^8 cells per ml was achieved.

Purification of OspA Protein from Borrelia Strain B31

Borrelia burgdorferi, either ^{14}C-palmitic acid-labeled or unlabeled, were harvested and washed as described (Brandt, M.E. *et al.*, Infect. Immun. 58:983-991 (1990)). Whole organisms were trypsinized according

to the protocol of Barbour *et al.* (Infect. Immun. 45:94-100 (1984)) with some modifications. The pellet was suspended in phosphate buffered saline (PBS, 10mM, pH 7.2), containing 0.8% tosyl-L-phenylalanine chloromethyl ketone (TPCK)-treated trypsin (Sigma, St. Louis, Missouri), the latter at a ratio of 1 μg per 10^8 cells. Reaction was carried out at 25°C for 1 hour, following which the cells were centrifuged. The pellet was washed in PBS with 100 μg/ml phenylmethylsulfonyl fluoride (PMSF). Triton X-114 partitioning of the pellet was carried out as described by Brandt *et al.* (Infect. Immun. 58:983-991 (1990)). Following trypsin treatment, cells were resuspended in ice-cold 2% (v/v) Triton X-114 in PBS at 10^9 cells per ml. The suspension was rotated overnight at 4°C, and the insoluble fraction removed as a pellet after centrifugation at 10,000 X g for 15 minutes at 4°C. The supernatant (soluble fraction) was incubated at 37°C for 15 minutes and centrifuged at room temperature at 1000 X g for 15 minutes to separate the aqueous and detergent phases. The aqueous phase was decanted, and ice cold PBS added to the lower Triton phase, mixed, warmed to 37°C, and again centrifuged at 1000 X g for 15 minutes. Washing was repeated twice more. Finally, detergent was removed from the preparation using a spin column of Bio-beads SM2 (BioRad, Melville, New York) as described (Holloway, P.W., Anal. Biochem. 53:304-308 (1973)).

Ion exchange chromatography was carried out as described by Dunn *et al.* (Prot. Exp. Purif. 1: 159-168 (1990)) with minor modifications. Crude OspA was dissolved in buffer A (1% Triton X-100, 10mM phosphate buffer (pH 5.0)) and loaded onto a SP Sepharose resin (Pharmacia, Piscataway, New Jersey), pre-equilibrated with buffer A at 25°C. After washing the column with 10

bed-volumes of buffer A, the bound OspA was eluted with buffer B (1% Triton X-100, 10mM phosphate buffer (pH 8.0)). OspA fractions were detected by protein assay using the BCA method (Pierce, Rockford, Illinois), or as radioactivity when intrinsically labeled material was fractionated. Triton X-100 was removed using a spin column of Bio-beads SM2.

This method purifies OspA from an outer surface membrane preparation. In the absence of trypsin-treatment, OspA and B were the major components of the soluble fraction obtained after Triton partitioning of strain B31. In contrast, when Triton extraction was carried out after trypsin-treatment, the OspB band is not seen. Further purification of OspA-B31 on a SP Sepharose column resulted in a single band by SDS-PAGE. The yield following removal of detergent was approximately 2 mg per liter of culture. This method of purification of OspA, as described herein for strain B31, can be used for other isolates of *Borrelia* as well. For strains such as strain K48, which lack OspB, trypsin treatment can be omitted.

Lipidation site of OspA-B31

^{14}C-palmitic acid labeled OspA from strain B31 was purified as described above and partially digested with endoproteinase Asp-N (data not shown). Following digestion, a new band of lower molecular weight was apparent by SDS-PAGE, found by direct amino-terminal sequencing to begin at Asp_{25}. This band had no trace of radioactivity by autoradiography (data not shown). OspA and B contain a signal sequence (L-X-Y-C) similar to the consensus described for lipoproteins of *E. coli*, and it has been predicted that the lipidation site of OspA and B should be the amino-terminal cysteine (Brandt, M.E. *et*

al., Infect. Immun 58: 983-991 (1990)). The results presented herein support this prediction.

B. Comparison of OspA Antibody Binding Regions in Nine Strains of *Borrelia burgdorferi*

The availability of the amino acid sequenced for OspA from a number of different isolates, combined with peptide mapping and Western blot analysis, permitted the identification of the antigenic domains recognized by monoclonal antibodies (MAbs) and allowed inference of the key amino acid residues responsible for specific antibody reactivity.

Strains of Borrelia burgdorferi

Nine strains of *Borrelia*, including seven European strains and two North American strains, were used in this study of antibody binding domains of several proteins. Information concerning the strains is summarized in Table I, below.

Table I. Representative *Borrelia* Strains

Strain	Location and Source	Reference for Strain
K48	Czechoslovakia, *Ixodes ricinus*	none
PGau	Germany, human ACA	Wilske, B. *et al.*, J. Clin. Microbiol. 32:340-350 (1993)
DK29	Denmark, human EM	Wilske, B. *et al.*
PKo	Germany, human EM	Wilske, B. *et al.*
PTrob	Germany, human skin	Wilske, B. *et al.*
Ip3	Khabarovsk, Russia, *I. persulcatus*	Asbrink, E. *et al.*, Acta Derm. Venereol. 64: 506-512 (1984)
Ip90	Khabarovsk, Russia, *I. persulcatus*	Asbrink, E. *et al.*
25015	Millbrook, NY, *I. persulcatus*	Barbour, A.G. *et al.*, Curr. Microbiol. 8:123-126 (1983)
B31	Shelter Island, NY, *I. scapularis*	Luft, B.J. *et al.*, Infect. Immun. 60: 4309-4321 (1992); ATCC 35210
PKa1	Germany, human CSF	Wilske, B. *et al.*
ZS7	Freiburg, Germany, *I. ricinus*	Wallich, R. *et al.*, Nucl. Acids Res. 17: 8864 (1989)
N40	Westchester Co., NY	Fikrig, E. *et al.*, Science 250:553-556 (1990)
PHei	Germany, human CSF	Wilske, B. *et al.*
ACAI	Sweden, human ACA	Luft, B. J. *et al.*, FEMS Microbiol. Lett. 93:73-68 (1992)
PBo	Germany, human CSF	Wilske, B. *et al.*

ACA = patient with acrodermatitis chronica atrophicans; EM = patient with erythema migrans; CSF = cerebrospinal fluid of patient with Lyme disease

Strains K48, PGau and DK29 were supplied by R. Johnson, University of Minnesota; PKo and pTrob were provided by B. Wilske and V. Preac-Mursic of the

Pettenkhofer Institute, Munich, Germany; and Ip3 and Ip90 were supplied by L. Mayer of the Center for Disease Control, Atlanta, Georgia. The North American strains included strain 25015, provided by J. Anderson of the Connecticut Department of Agriculture; and strain B31 (ATCC 35210).

Monoclonal Antibodies

Seven monoclonal antibodies (MAbs) were utilized in this study. Five of the MAbs (12, 13, 15, 83 and 336) were produced from hybridomas cloned and subcloned as previously described (Schubach, W.H., *et al.*, Infect. Immun. 59(6):1911-1915 (1991)). MAb H5332 (Barbour, A.G. *et al.*, Infect. Immun. 41:795-804 (1983)) was a gift from Drs. Alan Barbour, University of Texas, and MAb CIII.78 (Sears, J.E. *et al.*, J. Immunol. 147(6):1995-2000 (1991)) was a gift from Richard A. Flavell, Yale University. MAbs 12 and 15 were raised against whole sonicated B3; MAb 336 was produced against whole PGau; and MAbs 13 and 83 were raised to a truncated form of OspA cloned from the K48 strain and expressed in *E. coli* using the T7 RNA polymerase system (McGrath, B.C. *et al.*, Vaccines, Cold Spring Harbor Laboratory Press, Plainview, New York, pp. 365-370 (1993)). All MAbs were typed as being Immunoglobulin G (IgG).

Methods of Protein Cleavage, Western Blotting, and Amino-Terminal Sequencing

Prediction of the various cleavage sites was achieved by knowledge of the primary amino acid sequence derived from the full nucleotide sequences of OspA, many of which are currently available (see Table II, below). Cleavage sites can also be predicted based on the peptide sequence of OspA, which can be determined by standard techniques after isolation and purification of OspA by the method described above. Cleavage of several OspA isolates was

conducted to determine the localization of monoclonal antibody binding of the proteins.

Hydroxylamine-HCl (HA), N-chlorosuccinimide (NCS), and cyanogen bromide cleavage of OspA followed the methods described by Bornstein (Biochem. 9 (12):2408-2421 (1970)), Shechter *et al.*, (Biochem. 15 (23):5071-5075 (1976)), and Gross (in Hirs, C.H.W. (ed): Methods in Enzymology, (N.Y. Acad. Press), 11:238-255 (1967)) respectively. Protease cleavage by endoproteinase, Asp-N (Boehringer Mannheim, Indianapolis, Indiana), was performed as described by Cleveland D.W. *et al.*, (J. Biol. Chem. 252:1102-1106 (1977)). Ten micrograms of OspA were used for each reaction. The ratio of enzyme to OspA was approximately 1 to 10 (w/w).

Proteins and peptides generated by cleavage were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, U.K., Nature (London) 227:680-685 (1970)), and electroblotted onto immobilon Polyvinylidine Difluoride (PVDF) membranes (Ploskal, M.G. *et al.*, Biotechniques 4:272-283 (1986)). They were detected by amido black staining or by immunostaining with murine MAbs, followed by alkaline phosphatase-conjugated goat antimouse IgG. Specific binding was detected using a 5-bromo-4-chloro-3-indolylphosphate (BCIP)/nitroblue tetrazolium (NBT) developer system (KPL Inc., Gathersburg, Maryland).

In addition, amino-terminal amino acid sequence analysis was carried out on several cleavage products, as described by Luft *et al.* (Infect. Immun. 57:3637-3645 (1989)). Amido black stained bands were excised from PVDF blots and sequenced by Edman degradation using a Biosystems model 475A sequenator with model 120A PTH analyzer and model 900A control/data analyzer.

Cleavage Products of Outer Surface Protein A Isolates

Purified OspA-B31, labeled with ^{14}C-palmitic acid, was fragmented with hydroxylamine-HCl (HA) into two peptides, designated HA1 and HA2 (data not shown). The HA1 band migrated at 27 KD and retained its radioactivity, indicating that the peptide included the lipidation site at the N-terminus of the molecule (data not shown). From the predicted cleavage point, HA1 should correspond to residues 1 to 251 of OspA-B31. HA2 had a MW of 21.6 KD by SDS-PAGE, with amino-terminal sequence analysis showing it to begin at Gly72, i.e. residues 72 to 273 of OspA-B31. By contrast, HA cleaved OspA-K48 into three peptides, designated HA1, HA2, and HA3 with apparent MWs of 22KD, 16 KD and 12 KD, respectively. Amino-terminal sequencing showed HA1 to start at Gly72, and HA3 at Gly142. HA2 was found to have a blocked amino-terminus, as was observed for the full-length OspA protein. HA1, 2 and 3 of OspA-K48 were predicted to be residues 72-274, 1 to 141 and 142 to 274, respectively.

N-Chlorosuccinimide (NCS) cleaves tryptophan (W), which is at residue 216 of OspA-B31 or residue 217 of OspA-K48 (data not shown). NCS cleaved OspA-B31 into 2 fragments, NCS1, with MW of 23 KD, residues 1-216 of the protein, and NCS2 with a MW of 6.2 KD, residues 217 to 273 (data not shown). Similarly, K48 OspA was divided into 2 pieces, NCS1 residues 1-217, and NCS2 residues 218 to 274 (data not shown).

Cleavage of OspA by cyanogen bromide (CNBr) occurs at the carboxy side of methionine, residue 39. The major fragment, CNBr1, has a MW of 25.7 KD, residues 39-274 by amino-terminal amino acid sequence analysis (data not shown). CNBr2 (about 4 KD) could not be visualized by amido black staining; instead, lightly stained bands of about 20 KD MW were seen. These bands reacted with anti-

OspA MAbs, and most likely were degradation products due to cleavage by formic acid.

Determination of Antibody Binding Domains for Anti-OspA Monoclonal Antibodies

The cleavage products of OspA-B31 and OspA-K48 were analyzed by Western blot to assess their ability to bind to the six different MAbs. Preliminary Western blot analysis of the cleavage products demonstrated that strains K48 and DK29 have similar patterns of reactivity, as do IP3, PGau and PKo. The OspA of strain PTrob was immunologically distinct from the others, being recognized only by MAb 336. MAb 12 recognized only the two North American strains, B31 and 25015. When the isolates were separated into genogroups, it was remarkable that all the MAbs, except MAb 12, crossed over to react with multiple genogroups.

MAb12, specific for OspA-B31, bound to both HA1 and HA2 of OspA-B31. However, cleavage of OspA-B31 by NCS at residue Trp216 created fragments which did not react with MAb12, suggesting that the relevant domain is near or is structurally dependent upon the integrity of this residue (data not shown). MAb 13 bound only to OspA-K48, and to peptides containing the amino-terminus of that molecule (e.g. HA2; NCS1). It did not bind to CNBr1 residues 39 to 274. Thus the domain recognized by MAb13 is in the amino-terminal end of OspA-K48, near Met38.

MAb15 reacts with the OspA of both the B31 and K48 strains, and to peptides containing the N-terminus of OspA, such as HA1 of OspA-B31 and NCS1, but not to peptides HA2 of OspA-B31 and HA1 of OspA-K48 (data not shown). Both peptides include residue 72 to the C-terminus of the molecules. MAb15 bound to CNBr1 of OspA-K48, indicating the domain for this antibody to be residues 39 to 72, specifically near Gly72 (data not shown).

MAb83 binds to OspA-K48, and to peptides containing the C-terminal portion of the molecule, such as HA1. They do not bind to HA2 of OspA-K48, most likely because the C-terminus of HA2 of OspA-K48 ends at 141. Similar to MAb12 and OspA-B31, binding of MAbs 83 and CIII.78 is eliminated by cleavage of OspA at the tryptophan residue. Thus binding of MAbs 12, 83 and CIII.78 to OspA depends on the structural integrity of the Trp_{216} residue, which appears to be critical for antigenicity. Also apparent is that, although these MAbs bind to a common antigenic domain, the precise epitopes which they recognize are distinct from one another given the varying degrees of cross-reactivity to these MAbs among strains.

Although there is similar loss of binding activity of MAb336 with cleavage at Trp_{216}, this MAb does not bind to HA1 of OspA-B31, suggesting the domain for this antibody includes the carboxy-terminal end of the molecule, inclusive of residues 251 to 273. Low MW peptides, such as HA3 (10 KD) and NCS2 (6KD), of OspA-K48 do not bind this MAb on Western blots. In order to confirm this observation, we tested binding of the 6 MAbs with a recombinant fusion construct p3A/EC that contains a trpE leader protein fused with residues 217 to 273 of OspA-B31 (Schubach, W.H. *et al.*, Infect. Immun. 59(6): 1911-1915 (1991)). Only MAb336 reacted with this construct (data not shown). Peptides and antigenic domains localized by fragmentation of OspA are summarized in Figure 1.

Mapping of Domains to Define the Molecular Basis for the Serotype Analysis

To define the molecular basis for the serotype analysis of OspA, we compared the derived amino acid sequences of OspA for the nine isolates (Figure 2). At the amino terminus of the protein, these predictions can be more precise given the relatively small number of amino

acid substitutions in this region compared to the carboxy terminus. Domain 1, which is recognized by MAb13, includes residues Leu34 to Leu41. MAb13 only binds to the OspA of species K48, DK29 and IP90. Within this region, residue 37 is variable, however Gly37 is conserved amongst the three reactive strains. When Gly37 is changed to Glu37, as it is in OspA of strains B31, pTrob, PGau, and PKo, MAb13 does not recognize the protein (data not shown). By similar analysis, it can be seen that Asp70 is a crucial residue for Domain 2, which includes residues 65 to 75 and is recognized by MAb15. Domain 3 is reactive with MAbs H5332, 12 and 83, and includes residues 190-220. It is clear that significant heterogeneity exists between MAbs reactive with this domain, and that more than one conformational epitope must be contained within the sequence. Domain 4 binds MAb336, and includes residues 250 to 270. In this region, residue 266 is variable and therefore may be an important determinant. It is apparent, however, that other determinants of the reactivity of this monoclonal antibody reside in the region comprising amino acids 217-250. Furthermore, the structural integrity of Trp216 is essential for antibody reactivity in the intact protein. Finally, it is important to stress that Figure 2 indicates only the locations of the domains, and does not necessarily encompass the entire domain. Exact epitopes are being analyzed by site-directed mutagenesis of specific residues.

Overall, evidence suggests that the N-terminal portion is not the immunodominant domain of OspA, possibly by virtue of its lipidation, and the putative function of the lipid moiety in anchoring the protein to the outer envelope. The C-terminal end is immunodominant and includes domains that account in part for structural heterogeneity (Wilske, B. *et al.*, Med. Microbiol. Immunol. 181: 191-207 (1992)), and may provide epitopes for antibody

neutralization (Sears, J.E. *et al.*, J. Immunol. 147(6): 1995-2000 (1991)), and relate to other activities, such as the induction of T-cell proliferation (Shanafel, M.M., *et al.*, J. Immunol. 148: 218-224 (1992)). There are common epitopes in the carboxy-end of the protein that are shared among genospecies which may have immunoprotective potential (Wilske, B., *et al.*, Med. Microbiol. Immunol. 181: 191-207 (1992)).

Prediction of secondary structure on the basis of hydropathy analysis and circular dichroism and fluorescence spectroscopy measurements (McGrath, B.C., *et al.*, Vaccines, Cold Spring Harbor Laboratory Press, Plainview, New York; pp. 365-370 (1993)) suggest domains 3 and 4 to be in a region of the molecule with a propensity to form alpha-helix, whereas domains 1 and 2 occur in regions predicted to be beta-sheets (see Figure 1). These differences may distinguish domains in accessibility to antibody or to reactive T-cells (Shanafel, M.M. *et al.*, J. Immunol. 148: 218-224 (1992)). Site-directed mutagenesis of specific epitopes, as described below in Example 2, aids in identifying exact epitopes.

Example 2. Identification of an Immunologically Important Hypervariable Domain of the Major Outer Surface Protein A of *Borrelia*

This Example describes epitope mapping studies using chemically cleaved OspA and TrpE-OspA fusion proteins. The studies indicate a hypervariable region surrounding the single conserved tryptophan residue of OspA (at residue 216, or in some cases 217), as determined by a moving window population analysis of OspA from fifteen European and North American isolates of *Borrelia*. The hypervariable region is important for immune recognition.

Site-directed mutagenesis was also conducted to examine the hypervariable regions more closely. Fluorescence and circular dichroism spectroscopy have indicated that the conserved tryptophan is part of an alpha-helical region in which the tryptophan is buried in a hydrophobic environment (McGrath, B.C., *et al.*, Vaccines, Cold Spring Harbor Laboratory Press, Plainview, New York; pp. 365-370 (1993)). More polar amino acid side-chains flanking the tryptophan are likely to be exposed to the hydrophilic solvent. The hypervariability of these solvent-exposed residues among the various strains of *Borrelia* suggested that these amino acid residues may contribute to the antigenic variation in OspA. Therefore, site-directed mutagenesis was performed to replace some of the potentially exposed amino acid side chains in the protein from one strain with the analogous residues of a second strain. The altered proteins were then analyzed by Western Blot using monoclonal antibodies which bind OspA on the surface of the intact, non-mutated spirochete. The results indicated that certain specific amino acid changes near the tryptophan can abolish reactivity of OspA to these monoclonal antibodies.

A. Verification of Clustered Polymorphisms in Outer Surface Protein A Sequences

Cloning and sequencing of the OspA protein from fifteen European and North American isolates (described above in Table I) demonstrated that amino acid polymorphism is not randomly distributed throughout the protein; rather, polymorphism tended to be clustered in three regions of OspA. The analysis was carried out by plotting the moving, weighted average polymorphism of a window (a fixed length subsection of the total sequence) as it is slid along the sequence. The window size in this analysis was thirteen amino acids, based upon the determination of the largest

number of significantly deviating points as established by the method of Tajima (J. Mol. Evol. 33: 470-473 (1991)). The average weighted polymorphism was calculated by summing the number of variant alleles for each site. Polymorphism calculations were weighted by the severity of amino acid replacement (Dayhoff, M.O. *et al.*, in: Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure NBRF, Washington, Vol. 5, Suppl. 3: 345 (1978)). The sum was normalized by the window size and plotted. The amino acid sequence position corresponds to a window that encompasses amino acids 1 through 13. Bootstrap resampling was used to generate 95% confidence intervals on the sliding window analysis. Since *Borrelia* has been shown to be clonal, the bootstrap analysis should give a reliable estimate of the expected variance out of polymorphism calculations. The bootstrap was iterated five hundred times at each position, and the mean was calculated from the sum of all positions. The clonal nature of *Borrelia* ensures that the stochastic variance that results from differing genealogical histories of the sequence positions (as would be expected if recombination were prevalent) will be minimized.

This test verified that the three regions around the observed peaks all have significant excesses of polymorphism. Excesses of polymorphism were observed in the regions including amino acid residues 132-145, residues 163-177, and residues 208-221 (Figure 3). An amino acid alignment between residues 200 and 220 for B31, K48 and the four site-directed mutants is shown in Figure 4. The amino acid 208-221 region includes the region of OspA which has been modeled as an oriented alpha-helix in which the single tryptophan residue at amino acid 216 is buried in a hydrophobic pocket, thereby exposing more polar amino acids to the solvent (Figure 5) (France, L.L., *et al.*, Biochem. Biophys. Acta 1120: 59 (1992)). These potentially solvent-exposed residues showed considerable variability among the

OspAs from various strains and may be an important component of OspA antigenic variation. For the purposes of generating chimeric proteins, the hypervariable domains of interest are Domain A, which includes amino acid residues 120-140 of OspA; Domain B, which includes residues 150-180; and Domain C, which includes residues 200-216 or 217.

B. Site-Directed Mutagenesis of the Hypervariable Region

Site-directed mutagenesis was performed to convert residues within the 204-219 domain of the recombinant B31 OspA to the analogous residues of a European OspA variant, K48. In the region of OspA between residues 204 and 219, which includes the helical domain (amino acids 204-217), there are seven amino acid differences between OspA-B31 and OspA-K48. Three oligonucleotides were generated, each containing nucleotide changes which would incorporate K48 amino acids at their analogous positions in the B31 OspA protein. The oligos used to create the site-directed mutants were:

5'-CTTAATGACTCTGACACTAGTGC-3' (#613, which converts threonine at position 204 to serine, and serine at 206 to threonine (Thr204-Ser, Thr206-Ser)) (SEQ ID NO. 1);

5'-GCTACTAAAAAAACCGGGAAATGGAATTCA-3' (#625, which converts alanine at 214 to glycine, and alanine at 215 to lysine (Ala214-Gly, Ala215-Lys)) (SEQ ID NO. 2); and

5'-GCAGCTTGGGATTCAAAAACATCCACTTTAACA-3' (#640, which converts asparagine at 217 to aspartate, and glycine at 219 to lysine (Asn217-Asp, Gly219-Lys)) (SEQ ID NO. 3).

Site-directed mutagenesis was carried out by performing mutagenesis with pairs of the above oligos. Three site-directed mutants were created, each with two changes: OspA 613 (Thr204-Ser, Thr206-Ser), OspA 625 (Ala214-Gly, Ala215-Lys), and 640 (Asn217-Asp, Gly219-Lys). There were also two proteins with four changes: OspA 613/625 (Thr204-Ser, Thr206-Ser, Ala214-Gly, Ala215-Lys)

and OspA 613/640 (Thr204-Ser, Thr206-Ser, Asn217-Asp, Gly219-Lys).

Specificity of Antibody Binding to Epitopes of the Non-mutated Hypervariable Region

Monoclonal antibodies that agglutinate spirochetes, including several which are neutralizing *in vitro*, recognize epitopes that map to the hypervariable region around Trp216 (Barbour, A.G. *et al.*, Infect. and Immun. 41: 759 (1983); Schubach, W.H. *et al.*, Infect. and Immun. 59: 1911 (1991)). Western Blot analysis demonstrated that chemical cleavage of OspA from the B31 strain at Trp 216 abolishes reactivity of the protein with the agglutinating Mab 105, a monoclonal raised against B31 spirochetes (data not shown). The reagent, n-chlorosuccinimide (NCS), cleaves OspA at the Trp 216, forming a 23.2kd fragment and a 6.2kd peptide which is not retained on the Imobilon-P membrane after transfer. The uncleaved material binds Mab 105; however, the 23.2kd fragment is unreactive. Similar Western blots with a TrpE-OspA fusion protein containing the carboxy-terminal portion of the OspA protein demonstrated that the small 6.2kd piece also fails to bind Mab 105 (Schubach, W.H. *et al.*, Infect. and Immun. 59: 1911 (1991)).

Monoclonal antibodies H5332 and H3TS (Barbour, A.G. *et al.*, Infect. and Immun. 41: 759 (1983)) have been shown by immunofluorescence to decorate the surface of fixed spirochetes (Wilske, B. *et al.*, World J. Microbiol. 7: 130 (1991)). These monoclonals also inhibit the growth of the organism in culture. Epitope mapping with fusion proteins has confirmed that the epitopes which bind these Mabs are conformationally determined and reside in the carboxy half of the protein. Mab H5332 is cross-reactive among all of the known phylogenetic groups, whereas Mab H3TS and Mab 105 seem to be specific to the B31 strain to which they were

raised. Like Mab 105, the reactivities of H5332 and H3TS to OspA are abrogated by fragmentation of the protein at Trp216 (data not shown). Mab 336 was raised to whole spirochetes of the strain P/Gau. It cross-reacts to OspA from group 1 (the group to which B31 belongs) but not to group 2 (of which K48 is a member). Previous studies using fusion proteins and chemical cleavage have indicated that this antibody recognizes a domain of OspA in the region between residues 217 and 273 (data not shown). All of these Mabs will agglutinate the B31 spirochete.

Western Blot Analysis of Antibody Binding to Mutated Hypervariable Regions

Mabs were used for Western Blot analysis of the site-directed OspA mutants induced in *E.coli* using the T7 expression system (Dunn, J.J. *et al.*, Protein Expression and Purification 1: 159 (1990)). *E. coli* cells carrying Pet9c plasmids having a site-directed OspA mutant insert were induced at mid-log phase growth with IPTG for four hours at 37°C. Cell lysates were made by boiling an aliquot of the induced cultures in SDS gell loading dye, and this material was then loaded onto a 12% SDS gell (BioRad mini-Protean II), and electrophoresed. The proteins were then transferred to Imobilon-P membranes (Millipore) 70V, 2 hour at 4°C using the BioRad mini transfer system. Western analysis was carried out as described by Schubach *et al.* (Infect. Immun. 59: 1911 (1991)).

Western Blot analysis indicated that only the 625 mutant (Ala214-Gly and Ala215-Lys) retained binding to the agglutinating monoclonal H3TS (data not shown). However, the 613/625 mutant which has additional alterations to the amino terminus of Trp216 (Ser204-Thr and Thr206-Ser) did not bind this monoclonal. Both 640 and 613/640 OspAs which have the Asn217-Asp and Gly219-Lys changes on the carboxy-

terminal side of Trp216 also failed to bind Mab H3TS. This indicated that the epitope of the B31 OspA which binds H3TS is comprised of amino acid side-chains on both sides of Trp216.

The 613/625 mutant failed to bind Mabs 105 and H5332, while the other mutants retained their ability to bind these Mabs. This is important in light of the data using fusion proteins that indicate that Mab 105 behaves more like Mab H3TS in terms of its serotype specificity and binding to OspA (Wilske, B. *et al.*, Med. Microbiol. Immunol. 181: 191 (1992)). The 613/625 protein has, in addition to the differences at residues Thr204 and Ser206, changes immediately amino-terminal to Trp216 (Ala214-Gly and Ala215-Lys). The abrogation of reactivity of Mabs 105 and H5332 to this protein indicated that the epitopes of OspA which bind these monoclonals are comprised of residues on the amino-terminal side of Trp216.

The two proteins carrying the Asn217-Asp and Gly219-Lys replacements on the carboxy-terminal side of Trp216 (OspAs 640 and 613/640) retained binding to Mabs 105 and H5332; however, they failed to react with Mab 336, a monoclonal which has been mapped with TrpE-OspA fusion proteins and by chemical cleavage to a more carboxy-terminal domain. This result may explain why Mab 336 failed to recognize the K48-type of OspA (Group 2).

It is clear that amino acids Ser204 and Thr206 play an important part in the agglutinating epitopes in the region of the B31 OspA flanking Trp216. Replacement of these two residues altered the epitopes of OspA that bind Mabs 105, H3TS and H5332. The ability of the 640 changes alone to abolish reactivity of Mab 336 indicated that Thr204 and Ser206 are not involved in direct interaction with Mab 336.

The results indicated that the epitopes of OspA which are available to Mabs that agglutinate spirochetes are comprised at least in part by amino acids in the immediate

vicinity of Trp216. Since recent circular dichroism analysis indicated that the structures of B31 and K48 OspA differ very little within this domain, it is unlikely that the changes made by mutation have radically altered the overall structure of the OspA protein (France, L.L. *et al.*, Biochem. Biophys. Acta 1120: 59 (1992); and France *et al.*, Biochem. Biophys Acta, submitted (1993)). This hypothesis is supported by the finding that the recombinant, mutant OspAs exhibit the same high solubility and purification properties as the parent B31 protein (data not shown).

In summary, amino acid side-chains at Ser204 and Thr206 are important for many of the agglutinating epitopes. However, a limited set of conservative changes at these sites were not sufficient to abolish binding of all of the agglutinating Mabs. These results suggested that the agglutinating epitopes of OspA are distinct, yet may have some overlap. The results also supported the hypothesis that the surface-exposed epitope around Trp216 which is thought to be important for immune recognition and neutralization is a conformationally-determined and complex domain of OspA.

EXAMPLE 3. *Borrelia* Strains and Proteins

Proteins and genes from any strain of *Borrelia* can be utilized in the current invention. Representative strains are summarized in Table I, above.

A. Genes Encoding *Borrelia* Proteins

The chimeric peptides of the current invention can comprise peptides derived from any *Borrelia* proteins. Representative proteins include OspA, OspB, OspC, OspD, p12, p39, p41 (fla), p66, and p93. Nucleic acid sequences encoding several *Borrelia* proteins are presently available (see Table II, below); alternatively, nucleic acid

sequences encoding *Borrelia* proteins can be isolated and characterized using methods such as those described below.

Table II. References for Nucleic Acid Sequences for Several Proteins of Various *Borrelia* Strains

Strain	p93	OspA	p41 (fla)
K48	X69602 (SID 67)	X62624 (SID 8)	X69610 (SID 49)
PGau	SID 73	X62387 (SID 10)	X69612 (SID 51)
DK29	-	X63412 (SID 137)	X69608 (SID 53)
PKo	X69803 (SID 77)	X65599 (SID 141)	X69613 (SID 131)
PTrob	X69604 (SID 71)	X65598 (SID 135)	X69614 (SID 55)
Ip3	-	X70365 (SID 140)	-
Ip90	ND	Kryuchechnikov, V.N. *et al.*, J.Microbiol. Epid. Immunobiol. 12:41-44 (1988) (SID 138)	-
25015	X70365 (SID 75)	Fikrig, E.S. *et al.*, J. Immunol. 7:2256-2260 1992) SID 12)	-
B31	Perng, G.C. *et al.*, Infect. Immun. 59:2070-74 (1992); Luft, B.J. *et al.*, Infect. Immun. 60:4309-4321 (1992) (SID 65)	Bergstrom, S. *et al.*, Mol. Microbiol. 3:479-486 (1989) (SID 6)	Gassmann, G.S. *et al.*, Nucl. Acids Res. 17: 3590 (1989) (SID 127)
PKa1	-	X69606 (SID 132)	X69611 (SID 129)
ZS7	-	Jonsson, M. *et al.*, Infect. Immun. 60:1845-1853 (1992) (SID 134)	-
N40	-	Kryuchechnikov, V.N. *et al.* (SID 133)	-
PHei	-	X65600 (SID 136)	-
ACAI	-	Kryuchechnikov, V.N. *et al.* (SID 142)	-
PBo	X69601 (SID 69)	X65605 (SID 139)	X69610 (SID 130)

Numbers with an "X" prefix are GenBank data base accession numbers.
SID = SEQ ID NO.

B. Isolation of *Borrelia* Genes

Nucleic acid sequences encoding full length, lipidated proteins from known *Borrelia* strains were isolated using the polymerase chain reaction (PCR) as described below. In addition, nucleic acid sequences were generated which encoded truncated proteins (proteins in which the lipidation signal has been removed, such as by eliminating the nucleic acid sequence encoding the first 18 amino acids, resulting in non-lipidated proteins). Other proteins were generated which encoded polypeptides of a particular gene (i.e., encoding a segment of the protein which has a different number of amino acids than the protein does in nature). Using similar methods as those described below, primers can be generated from known nucleic acid sequences encoding *Borrelia* proteins and used to isolate other genes encoding *Borrelia* proteins. Primers can be designed to amplify all of a gene, as well as to amplify a nucleic acid sequence encoding truncated protein sequences, such as described below for OspC, or nucleic acid sequences encoding a polypeptide derived from a *Borrelia* protein. Primers can also be designed to incorporate unique restriction enzyme cleavage sites into the amplified nucleic acid sequences. Sequence analysis of the amplified nucleic acid sequences can then be performed using standard techniques.

Cloning and Sequencing of OspA Genes and Relevant Nucleic Acid Sequences

Borrelia OspA sequences were isolated in the following manner: 100 μl reaction mixtures containing 50 mM KCl, 10 mM TRIS-HCl (pH 8,3), 1.5 mM $MgCl_2$, 200 μM each NTP, 2.5 units of TaqI DNA polymerase (Amplitaq, Perkin-Elmer/Cetus) and 100 pmol each of the 5′ and 3′ primers (described below) were used. Amplification was performed in a Perkin-Elmer/Cetus thermal cycler as described (Schubach, W.H. *et*

al., Infect. Immun. 59:1811-1915 (1991)). The amplicon was visualized on an agarose gel by ethidium bromide staining. Twenty nanograms of the chloroform-extracted PCR product were cloned directly into the PC-TA vector (Invitrogen) by following the manufacturer's instructions. Recombinant colonies containing the amplified fragment were selected, the plasmids were prepared, and the nucleic acid sequence of each OspA was determined by the dideoxy chain-termination technique using the Sequenase kit (United States Biochemical). Directed sequencing was performed with M13 primers followed by OspA-specific primers derived from sequences, previously obtained with M13 primers.

Because the 5' and 3' ends of the OspA gene are highly conserved (Fikrig, E.S. *et al.*, J. Immunol. 7:2256-2260 (1992); Bergstrom, S. *et al.*, Mol. Microbiol. 3: 479-486 (1989); Zumstein, G. *et al.*, Med. Microbiol. Immunol. 181: 57-70 (1992)), the 5' and 3' primers for cloning can be based upon any known OspA sequences. For example, the following primers based upon the OspA nucleic acid sequence from strain B31 were used:

5'-GGAGAATATATTATGAAA-3' (-12 to +6) (SEQ ID NO. 4); and

5'-CTCCTTATTTTAAAGCG-3' (+826 to +809) (SEQ ID NO. 5).

(Schubach, W.H. *et al.*, Infect. Immun 59:1811-1915 (1991)).

OspA genes isolated in this manner include those for strains B31, K48, PGau, and 25015; the nucleic acid sequences are depicted in the sequence listing as SEQ ID NO. 6 (OspA-B31), SEQ ID NO. 8 (OspA-K48), SEQ ID NO. 10 (OspA-PGau), and SEQ ID NO. 12 (OspA-25015). An alignment of these and other OspA nucleic acid sequences is shown in Figure 42. The amino acid sequences of the proteins encoded by these nucleic acid sequences are represented as SEQ ID NO. 7 (OspA-B31), SEQ ID NO. 9 (OspA-K48), SEQ ID NO. 11 (OspA-PGau), and SEQ ID NO. 13 (OspA-25015).

The following primers were used to generate specific nucleic acid sequences of the OspA gene, to be used to

generate chimeric nucleic acid sequences (as described in Example 4):

5'-GTCTGCAAAAACCATGACAAG-3' (plus strand primer #369) (SEQ ID NO. 14);

5'-GTCATCAACAGAAGAAAAATTC-3' (plus strand primer #357) (SEQ ID NO 15);

5'-CCGGATCCATATGAAAAAATATTTATTGGG-3' (plus strand primer #607) (SEQ ID NO. 16);

5'-CCGGGATCCATATGGCTAAGCAAAATGTTAGC-3' (plus strand primer #584) (SEQ ID NO. 17);

5'-GCGTTCAAGTACTCCAGA-3' (minus strand primer #200) (SEQ ID NO. 18);

5'-GATATCTAGATCTTATTTTAAAGCGTT-3' (minus strand primer #586) (SEQ ID NO. 19); and

5'-GGATCCGGTGACCTTTTAAAGCGTTTTTAAT-3' (minus strand primer #1169) (SEQ ID NO. 20).

Cloning and Sequencing of OspB

Similar methods were also used to isolate OspB genes. One OspB genes isolated is represented as SEQ ID NO. 21 (OspB-B31); its encoded amino acid sequence is SEQ ID NO. 22.

The following primers were used to generate specific nucleic acid sequences of the OspB gene, to be used in generation of chimeric nucleic acid sequences (see Example 4):

5'-GGTACAATTACAGTACAA-3' (plus strand primer #721) (SEQ ID NO. 23);

5'-CCGAGAATCTCATATGGCACAAAAAGGTGCTGAGTCAATTGG-3' (plus strand primer #1105) (SEQ ID NO. 24);

5'-CCGATATCGGATCCTATTTTAAAGCGTTTTTAAGC-3' (minus strand primer # 1106) (SEQ ID NO. 25); and

5'-GGATCCGGTGACCTTTTAAAGCGTTTTTAAG-3' (minus strand primer #1170) (SEQ ID NO. 26).

Cloning and Sequencing of OspC

Similar methods were also used to isolate OspC genes. The following primers were used to isolate entire OspC genes from *Borrelia* strains B31, K48, PKO, and pTrob:

5'-GTGCGCGACCATATGAAAAAGAATACATTAAGTGCG-3' (plus strand primer having Nde1 site combined with start codon) (SEQ ID NO. 27), and

5'-GTCGGCGGATCCTTAAGGTTTTTTTGGACTTTCTGC-3' (minus strand primer having BamH1 site followed by stop codon) (SEQ ID NO. 28).

The nucleic acid sequences of the OspC genes were then determined by the dideoxy chain-termination technique using the Sequenase kit (United States Biochemical). OspC genes isolated and sequenced in this manner include those for strains B31, K48, PKo, and Tro; the nucleic acid sequences are depicted in the sequence listing as SEQ ID NO. 29 (OspC-B31), SEQ ID NO. 31 (OspC-K48), SEQ ID NO. 33 (OspC-PKo), and SEQ ID NO. 35 (OspC-Tro). An alignment of these sequences is shown in Figure 38. The amino acid sequences of the proteins encoded by these nucleic acid sequences are represented as SEQ ID NO. 30 (OspC-B31), SEQ ID NO. 32 (OspC-K48), SEQ ID NO. 34 (OspC-PKo), and SEQ ID NO. 36 (OspC-Tro).

Truncated OspC genes were generated using other primers. These primers were designed to amplify nucleic acid sequences, derived from the OspC gene, that lacked the nucleic acids encoding the signal peptidase sequence of the full-length protein. The primers corresponded to bp 58-75 of the natural protein, with a codon for Met-Ala attached ahead. For strain B31, the following primer was used:

5'-GTGCGCGACCATATGGCTAATAATTCAGGGAAAGAT-3' (SEQ ID NO. 37).

For strain PKo,

5'-GTGCGCGACCATATGGCTAGTAATTCAGGGAAAGGT-3' (SEQ ID NO. 38) was used.

For strains pTrob and K48,
5'-GTGCGCGACCATATGGCTAATAATTCAGGTGGGGAT-3' (SEQ ID NO. 39) was used.

Additional primers were also designed to amplify nucleic acids encoding particular polypeptides, for use in creation of chimeric nucleic acid sequences (see Example 4). These primers included:
5'-CTTGGAAAATTATTTGAA-3' (plus strand primer #520) (SEQ ID NO. 40);
5'-CACGGTCACCCCATGGGAAATAATTCAGGGAAAGG-3' (plus strand primer #58) (SEQ ID NO. 41);
5'-TATAGATGACAGCAACGC-3' (minus strand primer #207) (SEQ ID NO. 42); and
5'-CCGGTGACCCCATGGTACCAGGTTTTTTTGGACTTTCTGC-3' (minus strand primer #636) (SEQ ID NO. 43).

Cloning and Sequencing of OspD

Similar methods can be used to isolate OspD genes. An alignment of four OspD nucleic acid sequences (from strains pBo, PGau, DK29, and K48) is shown in Figure 39.

Cloning and Sequencing of p12

The p12 gene was similarly identified. Primers used to clone the entire p12 gene included: 5'-CCGGATCCATATGGTTAAAAAAATAATATTTATTTC-3' (forward primer # 757) (SEQ ID NO. 44); and 5'-GATATCTAGATCTTTAATTGCTCTGCTCACTCTCTTC-3' (reverse primer #758) (SEQ ID NO. 45).

To amplify a truncated p12 gene (one in which the transcribed protein is non-lipidated, and begins at amino acid 18 of the native sequence), the following primers were used: 5'-CCGGGATCCATATGGCTAGTGCAATTGGTCGTGG-3' (forward primer # 759) (SEQ ID NO. 46); and primer #758 (SEQ ID NO. 45).

Cloning and Sequencing of p41 (fla)

A similar approach was used to clone and sequence genes encoding the p41 (fla) protein. The p41 sequences listed in Table II with GenBank accession numbers were isolated using the following primers from strain B31:

5'-ATGATTATCAATCATAAT-3' (+1 to +18) (SEQ ID NO. 47); and

5'-TCTGAACAATGACAAAAC-3' (+1008 to +991) (SEQ ID NO. 48).

The nucleic acid sequences of p41 isolated in this manner are depicted in the sequence listing as SEQ ID NO. 51 (p41-PGau), and SEQ ID NO. 53 (p41-DK29). An alignment of several p41 nucleic acid sequences, including those for strains B31, pKa1, PGau, pBo, DK29, and pKo, is shown in Figure 41. The amino acid sequences of the proteins encoded by these nucleic acid sequences are represented as SEQ ID NO. 50 (p41-K48), SEQ ID NO. 52 (p41-PGau), SEQ ID NO. 54 (p41-DK29), SEQ ID NO. 56 (p41-PTrob), and SEQ ID NO. 58 (p41-PHei).

Other primers were designed to amplify nucleic acid sequences encoding polypeptides of p41, to be used in chimeric nucleic acid sequences. These primers included:

5'-TTGGATCCGGTCACCCCATGGCTCAATATAACCAATG-3' (minus strand primer #122) (SEQ ID NO. 59);

5'-TTGGATCCGGTCACCCCATGGCTTCTCAAAATGTAAG-3' (plus strand primer # 140) (SEQ ID NO. 60);

5'-TTGGATCCGGTGACCAACTCCGCCTTGAGAAGG-3' (minus strand primer # 234) (SEQ ID NO. 61); and

5'-TTGGATCCGGTGACCTATTTGAGCATAAGATGC-3' (minus strand primer #141) (SEQ ID NO. 62).

Cloning and Sequencing of p93

The same approach was also used to clone and sequence p93 protein. Genes encoding p93, as listed in Table II with GenBank accession numbers, were isolated by this method with the following primers from strain B31:

5'-GGTGAATTTAGTTGGTAAGG-3' (-54 to -35) (SEQ ID NO. 63); and

5'-CACCAGTTTCTTTAAGCTGCTCCTGC-3' (+1117 to +1092) (SEQ ID NO. 64).

The nucleic acid sequences of p93 isolated in this manner are depicted in the sequence listing as SEQ ID NO. 65 (p93-B31), SEQ ID NO. 67 (p93-K48) SEQ ID NO. 69 (p93-PBo), SEQ ID NO. 71 (p93-PTrob), SEQ ID NO. 73 (p93-PGau), SEQ ID NO. 75 (p93-25015), and SEQ ID NO. 77 (p93-PKo). The amino acid sequences of the proteins encoded by these nucleic acid sequences are represented as SEQ ID NO. 66 (p93-B31), SEQ ID NO. 68 (p93-K48) SEQ ID NO. 70 (p93-PBo), SEQ ID NO. 72 (p93-PTrob), SEQ ID NO. 74 (p93-PGau), SEQ ID NO. 76 (p93-25015), and SEQ ID NO. 78 (p93-PKo).

Other primers were used to amplify nucleic acid sequences encoding polypeptides of p93 to be used in generating chimeric nucleic acid sequences. These primers included:

5'-CCGGTCACCCCATGGCTGCTTTAAAGTCTTTA-3' (plus strand primer #475) (SEQ ID NO. 79);

5'-CCGGTCACCCCATGAATCTTGATAAAGCTCAG-3' (plus strand primer #900) (SEQ ID NO. 80);

5'-CCGGTCACCCCATGGATGAAAAGCTTTTAAAAAGT-3' (plus strand primer #1168) (SEQ ID NO. 81);

5'-CCGGTCACCCCCATGGTTGAGAAATTAGATAAG-3' (plus strand primer #1423) (SEQ ID NO. 82); and

5'-TTGGATCCGGTGACCCTTAACTTTTTTTAAAG-3' (minus strand primer # 2100) (SEQ ID NO. 83).

C. Expression of Proteins from Borrelia Genes

The nucleic acid sequences described above can be incorporated into expression plasmids, using standard techniques, and transfected into compatible host cells in order to express the proteins encoded by the nucleic acid

sequences. As an example, the expression the p12 gene and the isolation of p12 protein is set forth.

Amplification of the p12 nucleic acid sequence was conducted with primers that included a NdeI restriction site into the nucleic acid sequence. The PCR product was extracted with phenol/chloroform and precipitated with ethanol. The precipitated product was digested and ligated into an expression plasmid as follows: 15 μl (approximately 1 μg) of PCR DNA was combined with 2 μl 10X restriction buffer for NdeI (Gibco/BRL), 1 μl NdeI (Gibco/BRL), and 2 μl distilled water, and incubated overnight at 37°C. This mixture was subsequently combined with 3 μl 10X buffer (buffer 3, New England BioLabs), 1 μl BamHI (NEB), and 6 μl distilled water, and incubated at 37° for two hours. The resultant material was purified by preparative gel electrophoresis using low melting point agarose, and the band was visualized under long wave ultraviolet light and excised from the gel. The gel slice was treated with Gelase using conditions recommended by the manufacturer (Epicentre Technologies). The resulting DNA pelled was resuspended in 25-50 μl of 10 mM TRIS-CL (pH 8.0) and 1 mM EDTA (TE). An aliquot of this material was ligated into the Pet9c expression vector (Dunn, J. J. *et al.*, Protein Expression and Purification 1: 159 (1990)).

To ligate the material into the Pet9c expression vector, 20-50 ng of p12 nucleic acid sequences cut and purified as described above was combined with 5 μl 10 One-Phor-All (OPA) buffer (Pharmacia), 30-60 ng Pet9c cut with NdeI and BamHI, 2.5 μl 20 mM ATP, 2 μl T4 DNA ligase (Pharmacia) diluted 1:5 in 1X OPA buffer, and sufficient distilled water to bring the final volume to 50 μl. The mixture was incubated at 12°C overnight.

The resultant ligations were transformed into competent DH5-alpha cells and plated on nutrient agar plates containing 50 μg/ml kanamycin and incubated

overnight at 37 °C. DH5-alpha is used as a "storage strain" for T7 expression clones, because it is RecA deficient, so that recombination and concatenation are not problematic, and because it lacks the T7 RNA polymerase gene necessary to express the cloned gene. The use of this strain allows for cloning of potentially toxic gene products while minimizing the chance of deletion and/or rearrangement of the desired genes. Other cell lines having similar properties may also be used.

Kanamycin resistant colonies were single-colony purified on nutrient agar plates supplemented with kanamycin at 50 μg/ml. A colony from each isolate was inoculated into 3-5 ml of liquid medium containing 50 μg/ml kanamycin, and incubated at 37°C without agitation. Plasmid DNA was obtained from 1 ml of each isolate using a hot alkaline lysis procedure (Mantiatis, T. *et al.*, Molecular Cloning: A Laboratory Manual, cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982)).

Plasmid DNA was digested with EcoRI and BglII in the following manner: 15 μl plasmid DNA was combined with 2 μl 10X buffer 3 (NEB), 1 μ EcoRI (NEB), 1 μl BglII (NEB) and 1 μl distilled water, and incubated for two hours at 37°C. The entire reaction mixture was electrophoresed on an analytical agarose gel. Plasmids carrying the p12 insert were identified by the presence of a band corresponding to 925 base-pairs (full length p12) or 875 base-pairs (nonlipidated p12).

One or two plasmid DNAs from the full length and nonlipidated p12 clones in Pet9c were used to transform BL21 DE3 pLysS to kanamycin resistance as described by Studier *et al.* (Methods in Enzymology, Goeddel, D. (Ed.), Academic Press, 185: 60-89 (1990)). One or two transformants of the full length and nonlipidated clones were single-colony purified on nutrient plates containing 25 μg/ml chloramphenicol (to maintain pLysS) and 50 μg/ml

kanamycin at 37 °C. One colony of each isolate was inoculated into liquid medium supplemented with chloramphenicol and kanamycin and incubated overnight at 37°C. The overnight culture was subcultured the following morning into 500 ml of liquid broth with chloramphenicol (25 μg/ml) and kanamycin (50 μg/ml) and grown with aeration at 37°C in an orbital air-shaker until the absorbance at 600 nm reached 0.4-0.7. Isopropyl-thio-galactoside (IPTG) was added to a final concentration of 0.5 mM, for induction, and the culture was incubated for 3-4 hours at 37° as before. The induced cells were pelleted by centrifugation and resuspended in 25 ml of 20 mM $NaPO_4$ (pH 7.7). A small aliquot was removed for analysis by gel electrophoresis. Expressing clones produced proteins which migrated at the 12 kDa position.

A crude cell lysate was prepared from the culture as described for recombinant OspA by Dunn, J.J. *et al.*, (Protein Expression and Purification 1: 159 (1990)). The crude lysate was first passed over a Q-sepharose column (Pharmacia) which had been pre-equilibrated in Buffer A: 10 mM $NaPO_4$ (pH 7.7), 10 mM NaCl, 0.5 mM PMSF. The column was washed with 10 mM $NaPO_4$, 50 mM NaCl and 0.5 mM PMSF and then p12 was eluted in 10 mM $NaPO_4$, 0.5 mM PMSF with a NaCl gradient from 50-400 mM. p12 eluted approximately halfway through the gradient between 100 and 200 mM NaCl. The peak fractions were pooled and dialyzed against 10 mM NaPo4 (pH 7.7), 10 mM NaCl, 0.5 mM PMSF. The protein was then concentrated and applied to a Sephadex G50 gel filtration column of approximately 50 ml bed volume (Pharmacia), in 10 mM $NaPO_4$, 200 mM NaCl, 0.5 mM PMSF. p12 would typically elute shortly after the excluded volume marker. Peak fractions were determined by running small aliquots of all fractions on a gel. The p12 peak was pooled and stored in small aliquots at -20°C.

Example 4. Generation of Chimeric Nucleic Acid Sequences and Chimeric Proteins

A. General Protocol for Creation of Chimeric Nucleic Acid Sequences

The megaprimer method of site directed mutagenesis and its modification were used to generate chimeric nucleic acid sequences (Sarkar and Sommer, Biotechniques 8(4): 404-407 (1990); Aiyar, A. and J. Leis, Biotechniques 14(3): 366-369 (1993)). A 5′ primer for the first genomic template and a 3′ fusion oligo are used to amplify the desired region. the fusion primer consists of a 3′ end of the first template (DNA that encodes the amino-proximal polypeptide of the fusion protein), coupled to a 5′ end of the second template (DNA that encodes the carboxy-proximal polypeptide of the fusion protein).

The PCR amplifications are performed using Taq DNA polymerase, 10X PCR buffer, and $MgCl_2$ (Promega Corp., Madison, WI), and Ultrapure dNTPs (Pharmacia, Piscataway, NJ). One μg of genomic template 1, 5 μ of 10 μM 5′ oligo and 5 μl of 10 μM fusion oligo are combined with the following reagents at indicated final concentrations: 10X Buffer-Mg FREE (1X), $MgCl_2$ (2 mM), dNTP mix (200 μM each dNTP), Taq DNA polymerase (2.5 units), water to bring final volume to 100 μl. A Thermal Cycler (Perkin Elmer Cetus, Norwalk, CT) is used to amplify under the following conditions: 35 cycles at 95°C for one minute, 55°C for two minutes, and 72° for three minutes. This procedure results in a "megaprimer".

The resulting megaprimer is run on a 1X TAE, 4% low-melt agarose gel. The megaprimer band is cut from the gel and purified using the Promega Magic PCR Preps DNA purification system. Purified megaprimer is then used in a second PCR step. One μg of genomic template 2, approximately 0.5 μg of the megaprimer, and 5 μ of 10 μM 3′

oligo are added to a cocktail of 10X buffer, $MgCl_2$, dNTPs and Taq at the same final concentrations as noted above, and brought to 100 µl with water. PCR conditions are the same as above. The fusion product resulting from this amplification is also purified using the Promega Magic PCR Preps DNA purification system.

The fusion product is then ligated into TA vector and transformed into *E. coli* using the Invitrogen (San Diego, CA) TA Cloning Kit. Approximately 50 ng of PCR fusion product is ligated to 50 ng of pCRII vector with 1X Ligation Buffer, 4 units of T4 ligase, and brought to 10 Nl with water. This ligated product mixture is incubated at 12°C overnight (approximately 14 hours). Two µl of the ligation product mixture is added to 50 µl competent INC F' cells and 2 µ beta mercaptoethanol. The cells are then incubated for 30 minutes, followed by heat shock treatment at 42°C for 60 seconds, and an ice quenching for two minutes. 450 µl of warmed SOC media is then added to the cells, resulting in a transformed cell culture which is incubated at 37°C for one hour with slight shaking. 50 µl of the transformed cell culture is plated on LB + 50 µg/µl ampicillin plates and incubated overnight at 37°C. Single white colonies are picked and added to individual overnight cultures containing 3 ml LB with ampicillin (50 µg/µl).

The individual overnight cultures are prepared using Promega's Magic Miniprep DNA purification system. A small amount of the resulting DNA is cut using a restriction digest as a check. DNA sequencing is then performed to check the sequence of the fusion nucleic acid sequence, using the United States Biochemical (Cleveland, OH) Sequenase Version 2.0 DNA sequencing kit. Three to five µg of plasmid DNA is used per reaction. 2 µl 2M NaOH/2mM EDTA are added to the DNA, and the volume is brought to 20 µl with water. The mixture is then incubated at room temperature for five minutes. 7 µl water, 3µl 3M NaAc, 75

μl EtOH are added. The resultant mixture is mixed by vortex and incubated for ten minutes at -70°C, and then subjected to microfugation. After microfuge for ten minutes, the supernatant is aspirated off, and the pellet is dried in the speed vac for 30 second. 6 μl water, 2 μl annealing buffer, and 2 μl of 10 μM of the appropriate oligo is then added. This mixture is incubated for 10 minutes at 37°C and then allowed to stand at room temperature for 10 minutes. Subsequently, 5.5 μl of label cocktail (described above) is added to each sample of the mixture, which are incubated at room temperature for an additional five minutes. 3.5 μl labeled DNA is then added to each sample which is then incubated for five minutes at 37°C. 4 μl stop solution is added to each well. The DNA is denatured at 95° for two minutes, and then placed on ice.

Clones with the desired fusion nucleic acid sequences are then recloned in frame in the pEt expression system in the lipidated (full length) and non-lipidated (truncated, i.e., without first 17 amino acids) forms. The product is amplified using restriction sites contained in the PCR primers. The vector and product are cut with the same enzymes and ligated together with T4 ligase. The resultant plasmid is transformed into competent *E. coli* using standard transformation techniques. Colonies are screened as described earlier and positive clones are transformed into expression cells, such as *E. coli* BL21, for protein expression with IPTG for induction. The expressed protein in its bacterial culture lysate form and/or purified form is then injected in mice for antibody production. The mice are bled, and the sera collected for agglutination, *in vitro* growth inhibition, and complement- dependent and -independent lysis tests.

B. Specific Chimeric Nucleic Acid Sequences

Various chimeric nucleic acid sequences were generated. The nucleic acid sequences are described as encoding polypeptides from *Borrelia* proteins. The chimeric nucleic acid sequences are produced such that the nucleic acid sequence encoding one polypeptide is in the same reading frame as the nucleic acid sequence encoding the next polypeptide in the chimeric protein sequence encoded by the chimeric nucleic acid sequence. The proteins are listed sequentially (in order of presence of the encoding sequence) in the description of the chimeric nucleic acid sequence. For example, if a chimeric nucleic acid sequence consists of bp 1-650 from OspA-1 and bp 651-820 from OspA-2 were sequenced, the sequence of the chimer would include the first 650 base pairs from OspA-1 followed immediately by base pairs 651-820 of OspA-2.

OspA-K48/OspA-PGau A chimer of OspA from strain K48 (OspA-K48) and OspA from strain PGau (OspA-PGau) was generated using the method described above. This chimeric nucleic acid sequence included bp 1-654 from OspA-K48, followed by bp 655-820 from OspA-PGau. Primers used included: the amino-terminal sequence of OspA primer #607 (SEQ ID NO. 16); the fusion primer, 5'-AAAGTAGAAGTTTTTGAATCCCATTTTCCAGTTTTTTT-3' (minus strand primer #668-654) (SEQ ID NO. 84); the carboxy-terminal sequence of OspA primer #586 (SEQ ID NO. 19); and the sequence primers #369 (SEQ ID NO. 14) and #357 (SEQ ID NO. 15). The chimeric nucleic acid sequence is presented as SEQ ID NO. 85; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 86.

OspA-B31/OspA-PGau A chimer of OspA from strain B31 (OspA-B31) and OspA from strain PGau (OspA-PGau) was generated

using the method described above. This chimeric nucleic acid sequence included bp 1-651 from OspA-B31, followed by bp 652-820 from OspA-PGau. Primers used included: the fusion primer, 5'-AAAGTAGAAGTTTTTGAATTCCAAGCTGCAGTTTT-3' (minus strand primer #668-651) (SEQ ID NO. 87); and the sequence primer, #369 (SEQ ID NO. 14). The chimeric nucleic acid sequence is presented as SEQ ID NO. 88; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 89.

OspA-B31/OspA-K48 A chimer of OspA from strain B31 (OspA-B31) and OspA from strain K48 (OspA-K48) was generated using the method described above. This chimeric nucleic acid sequence included bp 1-651 from OspA-B31, followed by bp 652-820 from OspA-K48. Primers used included: the fusion primer, 5'-AAAGTGGAAGTTTTTGAATTCCAAGCTGCAGTTTTTTT-3' (minus strand primer #671-651) (SEQ ID NO. 90); and the sequence primer, #369 (SEQ ID NO. 14). The chimeric nucleic acid sequence is presented as SEQ ID NO. 91; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 92.

OspA-B31/OspA-25015 A chimer of OspA from strain B31 (OspA-B31) and OspA from strain 25015 (OspA-25015) was generated using the method described above. This chimeric nucleic acid sequence included bp 1-651 from OspA-B31, followed by bp 652-820 from OspA-25015. Primers used included: the fusion primer, 5'-TAAAGTTGAAGTGCCTGCATTCCAAGCTGCAGTTT-3' (SEQ ID NO. 93). The chimeric nucleic acid sequence is presented as SEQ ID NO. 94; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 95.

OspA-K48/OspA-B31/OspA-K48 A chimer of OspA from strain B31 (OspA-B31) and OspA from strain K48 (OspA-K48) was generated using the method described above. This chimeric nucleic acid sequence included bp 1-570 from OspA-B31, followed by bp 570-651 from OspA-B31, followed by bp 650-820 from OspA-K48. Primers used included: the fusion primer, 5'-CCCCAGATTTTGAAATCTTGCTTAAAACAAC-3' (SEQ ID NO. 96); and the sequence primer, #357 (SEQ ID NO. 15). The chimeric nucleic acid sequence is presented as SEQ ID NO. 97; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 98.

OspA-B31/OspA-K48/OspA-B31/OspA-K48 A chimer of OspA from strain B31 (OspA-B31) and OspA from strain K48 (OspA-K48) was generated using the method described above. This chimeric nucleic acid sequence included bp 1-420 from OspA-B31, followed by 420-570 from OspA-K48, followed by bp 570-650 from OspA-B31, followed by bp 651-820 from OspA-K48. Primers used included: the fusion primer, 5'-CAAGTCTGGTTCCAATTTGCTCTTGTTATTAT-3' (minus strand primer #436-420) (SEQ ID NO. 99); and the sequence primer, #357 (SEQ ID NO. 15). The chimeric nucleic acid sequence is presented as SEQ ID NO. 100; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 101.

OspA-B31/OspB-B31 A chimer of OspA and OspB from strain B31 (OspA-B31, OspB-B31) was generated using the method described above. The chimeric nucleic acid sequence included bp 1-651 from OspA-B31, followed by bp 652-820 from OspB-B31. Primers used included: the fusion primer, 5'-GTTAAAGTGCTAGTACTGTCATTCCAAGCTGCAGTTTTTTT-3' (minus strand primer #740-651) (SEQ ID NO. 102); the carboxy-terminal sequence of OspB primer #1106 (SEQ ID NO. 25); and the sequence primer #357 (SEQ ID NO. 15). The chimeric

nucleic acid sequence is presented as SEQ ID NO. 103; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 104.

OspA-B31/OspB-B31/OspC-B31 A chimer of OspA, OspB and OspC from strain B31 (OspA-B31, OspB-B31, and OspC-B31) was generated using the method described above. The chimeric nucleic acid sequence included bp 1-650 from OspA-B31, followed by bp 652-820 from OspB-B31, followed by bp 74-630 of OspC-B31. Primers used included: the fusion primer, 5'-TGCAGATGTAATCCCATCCGCCATTTTTAAAGCGTTTTT-3' (SEQ ID NO. 105); and the carboxy-terminal sequence of OspC primer (SEQ ID NO. 28). The chimeric nucleic acid sequence is presented as SEQ ID NO. 106; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 107.

OspC-B31/OspA-B31/OspB-B31 A chimer of OspA, OspB and OspC from strain B31 (OspA-B31, OspB-B31, and OspC-B31) was generated using the method described above. The chimeric nucleic acid sequence included bp 1-630 from OspC-B31, followed by bp 52-650 from OspA-B31, followed by bp 650-820 of OspB-B31. Primers used included: the amino-terminal sequence of OspC primer having SEQ ID NO. 27; the fusion primer, 5'-GCTGCTAACATTTTGCTTAGGTTTTTTTGGACTTTC-3' (minus strand primer #69-630) (SEQ ID NO. 108); and the sequence primers #520 (SEQ ID NO. 40) and #200 (SEQ ID NO. 18). The chimeric nucleic acid sequence is presented as SEQ ID NO. 109; the chimeric protein encoded by this chimeric nucleic acid sequence is presented as SEQ ID NO. 110.

Additional Chimeric Nucleic Acid Sequences

Using the methods described above, other chimeric nucleic acid sequences were produced. These chimeric

nucleic acid sequences, and the proteins encoded, are summarized in Table 3.

Table III Chimeric Nucleic acid Sequences and the Encoded Proteins

Chimers Generated (base pairs)	SEQ ID NO. (nt)	SEQ ID NO. (protein)
OspA (52-882) / p93 (1168-2100)	111	112
OspB (45-891) / p41 (122-234)	113	114
OspB (45-891) / p41 (122-295)	115	116
OspB (45-891) / p41 (140-234)	117	118
OspB (45-891) / p41 (140-295)	119	120
OspB (45-891) / p41 (122-234) / OspC (58-633)	121	122
OspA-Tro/OspA-Bo	137	138
OspA-PGau/OspA-Bo	139	140
OspA-B31/OspA-PGau/OspA-B31/ OspA-K48	141	142
OspA-PGau/OspA-B31/OspA-K48	143	144

C. Purification of Proteins Generated by Chimeric Nucleic Acid Sequences

The chimeric nucleic acid sequences described above, as well as chimeric nucleic acid sequences produced by the methods described above, are used to produce chimeric proteins encoded by the nucleic acid sequences. Standard methods, such as those described above in Example 3, concerning the expression of proteins from *Borrelia* genes, can be used to express the proteins in a compatible host organism. The chimeric proteins can then be isolated and purified using standard techniques.

If the chimeric protein is soluble, it can be purified on a Sepharose column. Insoluble proteins can be solubilized in guanidine and purified on a Ni++ column;

alternatively, they can be solubilized in 10 mM $NaPO_4$ with 0.1 - 1% TRIXON X 114, and subsequently purified over an S column (Pharmacia). Lipidated proteins were generally purified by the latter method. Solubility was determined by separating both soluble and insoluble fractions of cell lysate on a 12% PAGE gel, and checking for the localization of the protein by Coomasie staining, or by Western blotting with monoclonal antibodies directed to an antigenic polypeptide of the chimeric protein.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. such equivalents are intended to be encompassed in the scope of the following claims.

CLAIMS

What is claimed is:

1. A chimeric protein comprising two or more antigenic *Borrelia* polypeptides, wherein the antigenic *Borrelia* polypeptides which comprise the chimeric protein do not occur naturally in the same protein in *Borrelia*.

2. The chimeric protein of Claim 1, wherein the antigenic *Borrelia* polypeptides are from two or more different species of *Borrelia*.

3. The chimeric protein of Claim 2, wherein the antigenic *Borrelia* polypeptides are derived from *Borrelia* proteins selected from the group consisting of: outer surface protein A, outer surface protein B, outer surface protein C, outer surface protein D, p12, p39, p41, p66, and p93.

4. The chimeric protein of Claim 3, wherein the antigenic *Borrelia* polypeptides are from corresponding proteins from two or more different species of *Borrelia*.

5. The chimeric protein of Claim 3, wherein the antigenic *Borrelia* polypeptides are from non-corresponding proteins from at least two different species of *Borrelia*.

6. The chimeric protein of Claim 1, wherein two or more antigenic *Borrelia* polypeptides are from the same species of *Borrelia*.

7. The chimeric protein of Claim 6, wherein the antigenic *Borrelia* polypeptides are derived from *Borrelia* proteins selected from the group consisting of: outer surface protein A, outer surface protein B, outer surface protein C, outer surface protein D, p12, p39, p41, p66, and p93.

8. The chimeric protein of Claim 7, wherein the antigenic *Borrelia* polypeptides are from the same protein.

9. The chimeric protein of Claim 6, wherein the antigenic *Borrelia* polypeptides are from different proteins.

10. A chimeric protein comprising two antigenic *Borrelia* polypeptides flanking a tryptophan residue, wherein the amino-proximal polypeptide consists of a polypeptide that is proximal from the single tryptophan residue of a first outer surface protein of *Borrelia*, and the carboxy-proximal polypeptide consists of a polypeptide that is distal from the single tryptophan residue of a second outer surface protein of *Borrelia*.

11. The chimeric protein of Claim 10, wherein the first and second outer surface proteins are from the same species of *Borrelia*.

12. The chimeric protein of Claim 11, wherein the first outer surface protein is outer surface protein A and the second outer surface protein is outer surface protein B.

13. The chimeric protein of Claim 11, wherein the first outer surface protein is outer surface protein B, and

the second outer surface protein is outer surface protein A.

14. The chimeric protein of Claim 10, wherein the first and second outer surface proteins are from different species of *Borrelia*.

15. The chimeric protein of Claim 14, wherein the first outer surface protein is outer surface protein A and the second outer surface protein is outer surface protein B.

16. The chimeric protein of Claim 14, wherein the first outer surface protein is outer surface protein B, and the second outer surface protein is outer surface protein A.

17. The chimeric protein of Claim 14, wherein the first and second outer surface proteins are corresponding proteins selected from the group consisting of: outer surface protein A and outer surface protein B.

18. The chimeric protein of Claim 10, wherein the first outer surface protein is outer surface protein A and the second outer surface protein is outer surface protein B.

19. The chimeric protein of Claim 18, wherein the amino-proximal polypeptide further comprises a first, second, and third hypervariable domain, the first hypervariable domain consisting of residues 120 through 140 of outer surface protein A, the second hypervariable domain consisting of residues 150 through 180 of outer surface protein A, and the third

hypervariable domain consisting of residues 200 through 217 of outer surface protein A.

20. The chimeric protein of Claim 19, wherein the first and second hypervariable domains are derived from outer surface protein A from different species of *Borrelia*.

21. The chimeric protein of Claim 10, further comprising an antigenic *Borrelia* polypeptide derived from a *Borrelia* protein selected from the group consisting of: outer surface protein A, outer surface protein B, outer surface protein C, outer surface protein D, p12, p39, p41, p66, and p93.

22. A nucleic acid sequence encoding a chimeric protein comprising two antigenic *Borrelia* polypeptides, wherein the two antigenic *Borrelia* polypeptides which comprise the chimeric protein do not occur naturally in the same protein in *Borrelia*.

23. The nucleic acid sequence of Claim 22, wherein the antigenic *Borrelia* polypeptides are from two or more different species of *Borrelia*.

24. The nucleic acid sequence of Claim 23, wherein the antigenic *Borrelia* polypeptides are derived from *Borrelia* proteins selected from the group consisting of: outer surface protein A, outer surface protein B, outer surface protein C, outer surface protein D, p12, p39, p41, p66, and p93.

25. The nucleic acid sequence of Claim 24, wherein the antigenic *Borrelia* polypeptides are from corresponding

proteins from two or more different species of *Borrelia*.

26. The nucleic acid sequence of Claim 24, wherein two or more of the antigenic *Borrelia* polypeptides are from non-corresponding proteins from different species of *Borrelia*.

27. The nucleic acid sequence of Claim 22, wherein two or more antigenic *Borrelia* polypeptides are from the same species of *Borrelia*.

28. The nucleic acid sequence of Claim 27, wherein the antigenic *Borrelia* polypeptides are derived from *Borrelia* proteins selected from the group consisting of: outer surface protein A, outer surface protein B, outer surface protein C, outer surface protein D, p12, p39, p41, p66, and p93.

29. The nucleic acid sequence of Claim 28, wherein the antigenic *Borrelia* polypeptides are from the same protein.

30. The nucleic acid sequence of Claim 27, wherein the antigenic *Borrelia* polypeptides are from different proteins.

31. A nucleic acid sequence encoding a chimeric protein comprising two antigenic *Borrelia* polypeptides flanking a tryptophan residue, wherein the amino-proximal polypeptide consists of a polypeptide that is proximal from the single tryptophan residue of a first outer surface protein of *Borrelia*, and the carboxy-proximal polypeptide consists of a polypeptide that is distal from the single tryptophan residue of a second outer surface protein of *Borrelia*.

32. The nucleic acid sequence of Claim 31, wherein the first and second outer surface proteins are from the same species of *Borrelia*.

33. The nucleic acid sequence of Claim 32, wherein the first outer surface protein is outer surface protein A and the second outer surface protein is outer surface protein B.

34. The nucleic acid sequence of Claim 32, wherein the first outer surface protein is outer surface protein B, and the second outer surface protein is outer surface protein A.

35. The nucleic acid sequence of Claim 31, wherein the first and second outer surface proteins are from different species of *Borrelia*.

36. The nucleic acid sequence of Claim 35, wherein the first outer surface protein is outer surface protein A and the second outer surface protein is outer surface protein B.

37. The nucleic acid sequence of Claim 35, wherein the first outer surface protein is outer surface protein B, and the second outer surface protein is outer surface protein A.

38. The nucleic acid sequence of Claim 35, wherein the first and second outer surface proteins are corresponding proteins selected from the group consisting of: outer surface protein A and outer surface protein B.

39. The nucleic acid sequence of Claim 31, wherein the first outer surface protein is outer surface protein A and the second outer surface protein is outer surface protein B.

40. The nucleic acid sequence of Claim 39, wherein the amino-proximal polypeptide further comprises a first and a second hypervariable domain, the first hypervariable domain consisting of amino acid residues 1 through 140 of outer surface protein A, and the second hypervariable domain consisting of amino acid residues 150 through 217 of outer surface protein A.

41. The nucleic acid sequence of Claim 40, wherein the first and second hypervariable domains are derived from outer surface protein A from different species of *Borrelia*.

42. The nucleic acid sequence of Claim 31, further comprising an antigenic *Borrelia* polypeptide derived from a *Borrelia* protein selected from the group consisting of: outer surface protein A, outer surface protein B, outer surface protein C, outer surface protein D, p12, p39, p41, p66, and p93.

43. A nucleic acid sequence having a sequence selected from the group consisting of: SEQ ID NO. 85, SEQ ID NO. 88, SEQ ID NO. 91, SEQ ID NO. 94, SEQ ID NO. 97, SEQ ID NO. 100, SEQ ID NO. 103, SEQ ID NO. 106, SEQ ID NO. 109, SEQ ID NO. 111, SEQ ID NO. 113, SEQ ID NO. 115, SEQ ID NO. 117, SEQ ID NO. 119, SEQ ID NO. 121, SEQ ID NO. 137, SEQ ID NO. 139, SEQ ID NO. 141, and SEQ ID NO. 143.

44. A protein having an amino acid sequence selected from the group consisting of: SEQ ID NO. 86, SEQ ID NO. 89, SEQ ID NO. 92, SEQ ID NO. 95, SEQ ID NO. 98, SEQ ID NO. 101, SEQ ID NO. 104, SEQ ID NO. 107, SEQ ID NO. 110, SEQ ID NO. 112, SEQ ID NO. 114, SEQ ID NO. 116, SEQ ID NO. 118, SEQ ID NO. 120, SEQ ID NO. 122, SEQ ID NO. 138, SEQ ID NO. 140, SEQ ID NO. 142, and SEQ ID NO. 144.

45. A chimeric protein according to any one of claims 1 to 21 and 44 for use in therapy or diagnosis, for example as a vaccine against Borrelia infection, in immunodiagnostic assays to detect the presence of antibodies to Borrelia or to measure T-cell reactivity.

46. A chimeric protein according to claim 45, wherein the immunodiagnostic assay is a dot blot, Western blot, ELISA or agglutination assay.

47. Use of the chimeric protein according to any one of claims 1 to 21 and 44, or the nucleic acid sequence of any one of claims 22 to 43, for the manufacture of a compound for use in therapy or diagnosis, for example as a vaccine against Borrelia infection, in immunodiagnostic assays to detect the presence of antibodies to Borrelia or to measure T-cell reactivity.

48. Use according to claim 47, wherein the immunodiagnostic assay is a dot blot, Western blot, ELISA or agglutination assay.

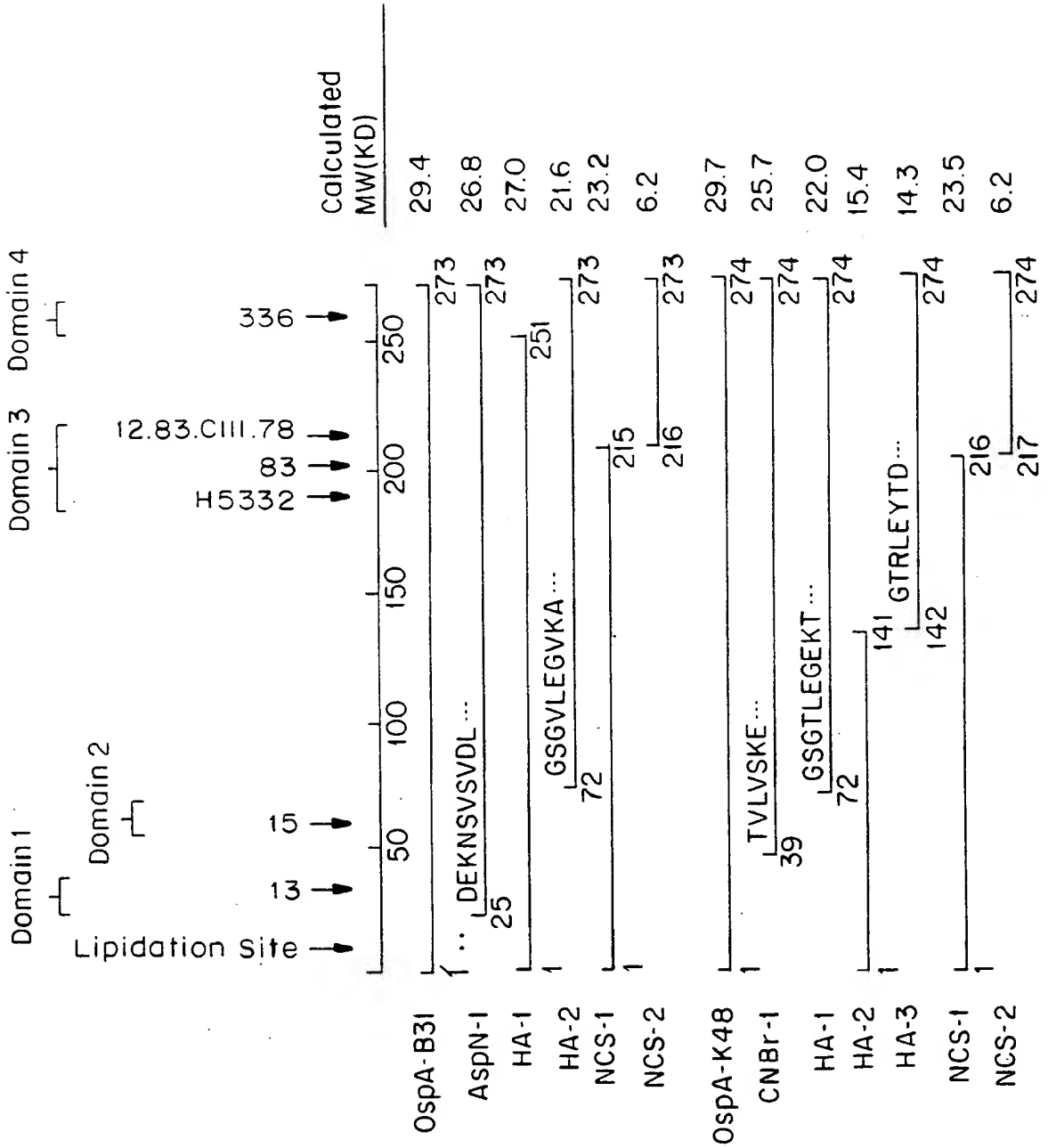

FIG. 1

Domain 1

	34	35	36	37	38	39	40	41
A-B31	L	P	G	E	M	K	V	L
A-TRo	L	P	G	E	M	K	V	L
A-K48	L	P	G	G	M	T	V	L
A-DK29	L	P	G	G	M	T	V	L
A-P/Gau	L	P	G	E	M	K	V	L
A-PKo	L	P	G	E	M	K	V	L
A-IP3	L	P	G	E	I	K	V	L
A-IP90	L	P	G	G	M	G	V	L
A-25015	L	P	G	E	M	K	V	L

Domain 2

	65	66	67	68	69	70	71	72	73	74	75
A-B31	G	T	S	D	K	N	N	G	S	G	V
A-TRo	G	T	S	D	K	S	N	G	S	G	T
A-K48	G	T	S	D	K	N	N	G	S	G	T
A-DK29	G	T	S	D	K	N	N	G	S	G	T
A-P/Gau	G	T	S	D	K	D	N	G	S	G	T
A-PKo	G	T	S	D	K	D	N	G	S	G	T
A-IP3	G	T	S	D	K	D	N	G	S	G	V
A-IP90	G	T	S	D	K	N	N	G	S	G	T
A-25015	G	T	S	D	K	N	N	G	S	G	V

Domain 3

	190	200	210	220
A-B31	NISKSGEVSVELNDTDSSAATKKTAAWNSGT			
A-TRo	HIPNSGEITVELNDSNSTQATKKTGKWDSNT			
A-K48	NILKSGEITVALDDSDTTQATKKTGKWDSKT			
A-DK29	NILKSGEITAALDDSDTTRATKKTGKWDSKT			
A-P/Gau	EIAKSGEVTVALNDTNTTQATKKTGAWDSKT			
A-PKo	EIAKSGEVTVALNDTNTTQATKKTGAWDSKT			
A-IP3	EIAKSGEVTVALNDTNTTQATKKTGAWDSKT			
A-IP90	HISNSGEITVELNDSDTTQATKKTGTWDSKT			
A-25015	HISKSGEVTAELNDTDSTQATKKTGKWDAGT			

Domain 4

	250	260	270
A-B31	SNGTKLEGSAVEITKLDEIKN		
A-TRo	SAGTNLEGNAVEIKTLDELKN		
A-K48	SAGTNLEGKAVEITTLKELKN		
A-DK29	SAGTNLEGKAVEITTLKELKN		
A-P/Gau	SAGTNLEGTAVEIKTLDELKN		
A-PKo	SAGTNLEGTAVEIKTLDELKN		
A-IP3	SAGTNLEGTAVEIKTLDELKN		
A-IP90	SAGTNLEGKAVEITTLKELKN		
A-25015	SAGTNLEGTAVEIKTLDEIKN		

Figure 2

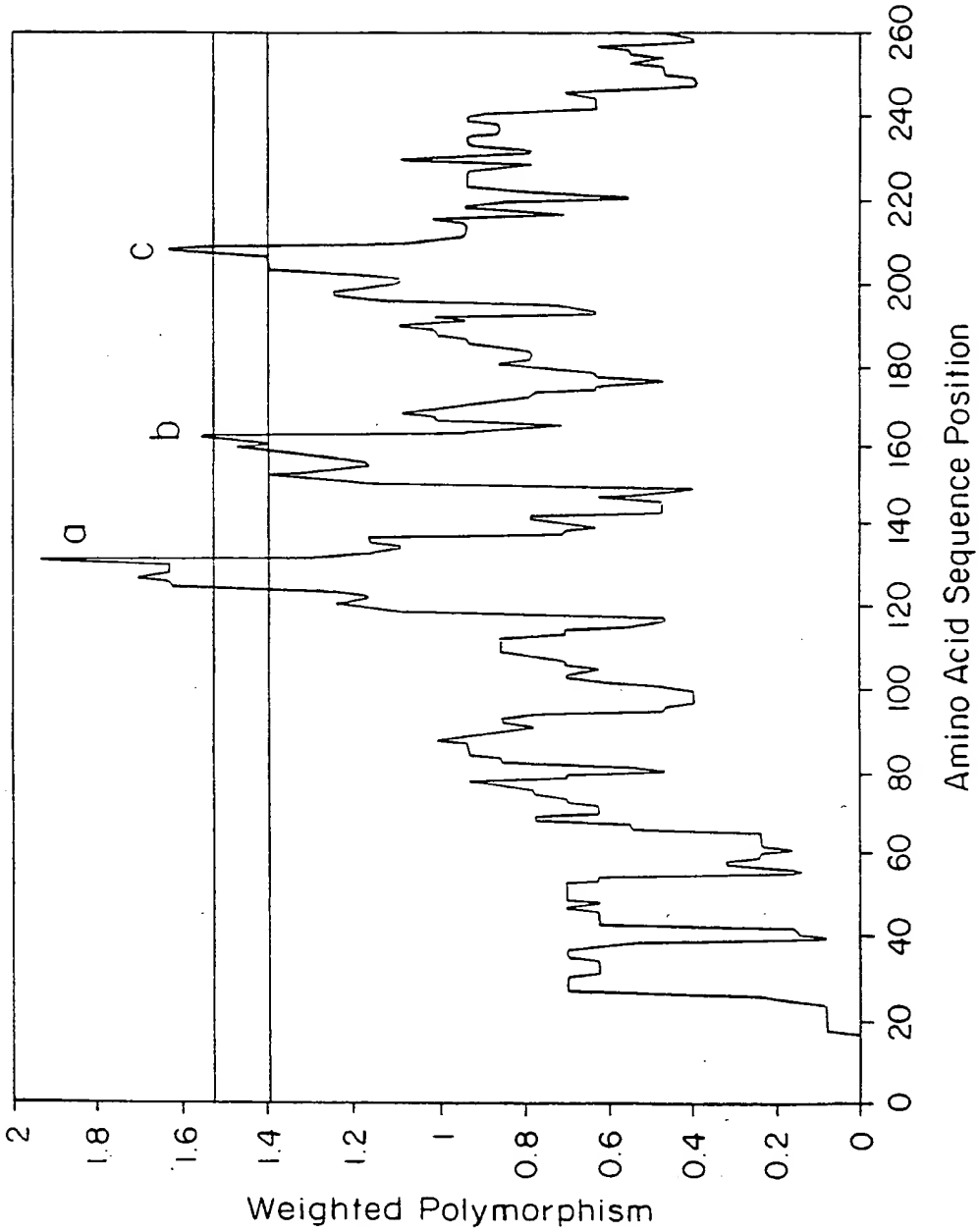

FIG. 3

```
                1
B31:    ELNDTDSSAATKKTAAWNSGT
K48:    ALDDSDTTQATKKTGKWDSKT

613:    ELNDSDISAATKKTAAWNSGT
625:    ELNDTDSSAATKKTGKWNSGT
640:    ELNDTDSSAATKKTAAWDSKT
613/625: ELNDSDISAATKKTGKWNSGT
613/640: ELNDSDISAATKKTAAWDSKT
```

Figure 4

5/133

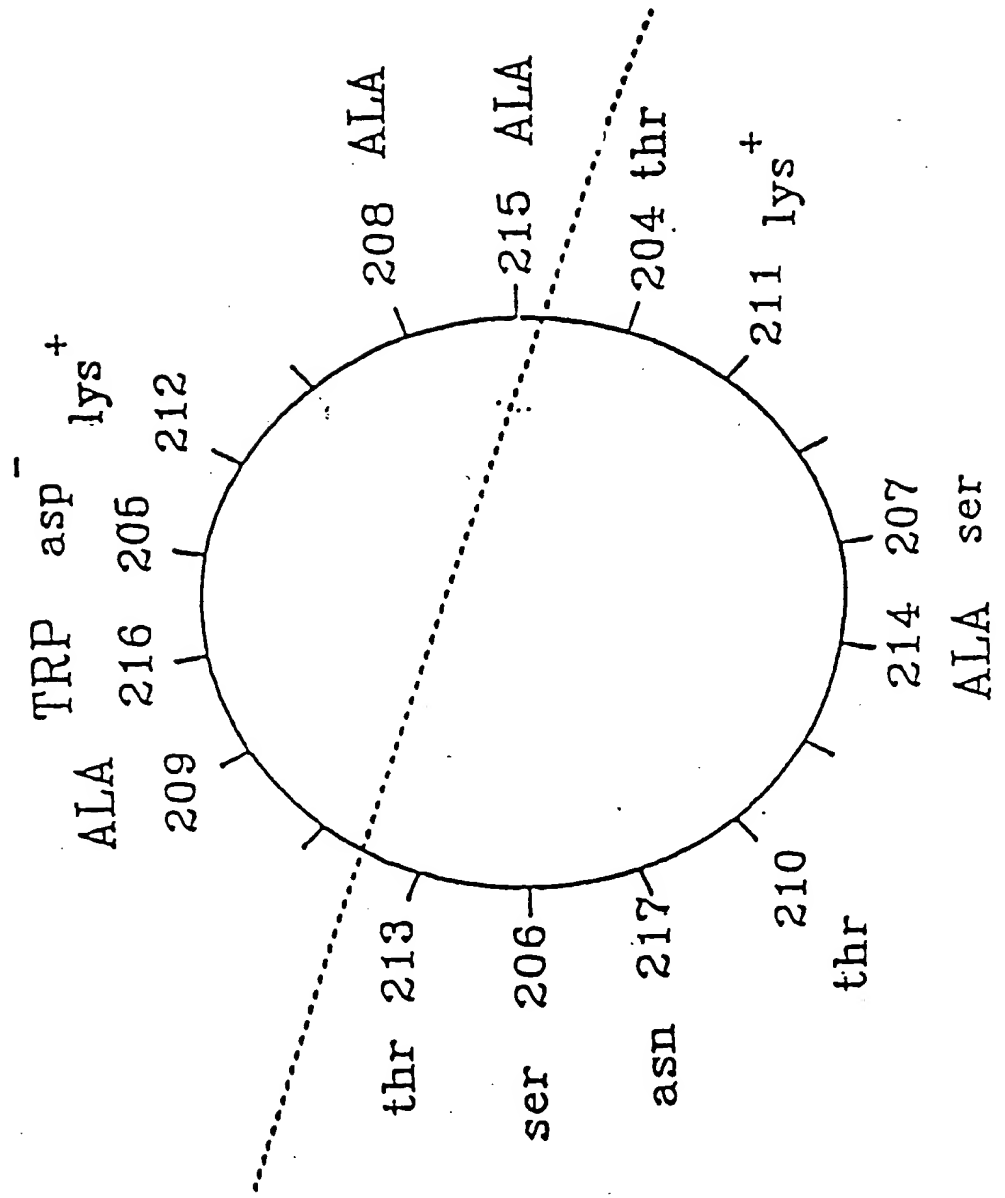

Figure 5

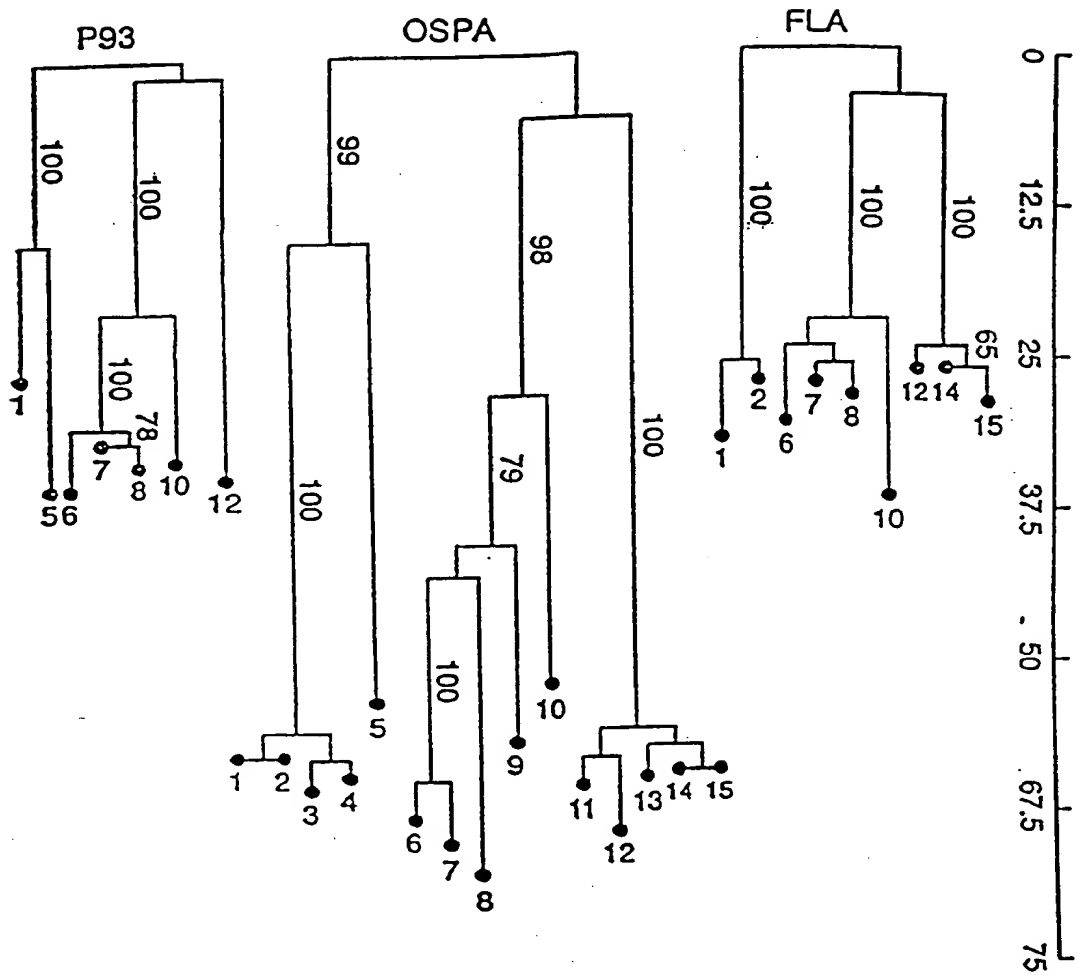

Figure 6

```
ATG AAA AAA TAT TTA TTG GGA ATA GGT CTA ATA TTA GCC TTA ATA GCA        48
Met Lys Lys Tyr Leu Leu Gly Ile Gly Leu Ile Leu Ala Leu Ile Ala
  1               5                  10                  15

TGT AAG CAA AAT GTT AGC AGC CTT GAC GAG AAA AAC AGC GTT TCA GTA        96
Cys Lys Gln Asn Val Ser Ser Leu Asp Glu Lys Asn Ser Val Ser Val
            20                  25                  30

GAT TTG CCT GGT GAA ATG AAA GTT CTT GTA AGC AAA GAA AAA AAC AAA       144
Asp Leu Pro Gly Glu Met Lys Val Leu Val Ser Lys Glu Lys Asn Lys
        35                  40                  45

GAC GGC AAG TAC GAT CTA ATT GCA ACA GTA GAC AAG CTT GAG CTT AAA       192
Asp Gly Lys Tyr Asp Leu Ile Ala Thr Val Asp Lys Leu Glu Leu Lys
    50                  55                  60

GGA ACT TCT GAT AAA AAC AAT GGA TCT GGA GTA CTT GAA GGC GTA AAA       240
Gly Thr Ser Asp Lys Asn Asn Gly Ser Gly Val Leu Glu Gly Val Lys
65                  70                  75                  80

GCT GAC AAA AGT AAA GTA AAA TTA ACA ATT TCT GAC GAT CTA GGT CAA       288
Ala Asp Lys Ser Lys Val Lys Leu Thr Ile Ser Asp Asp Leu Gly Gln
                85                  90                  95

ACC ACA CTT GAA GTT TTC AAA GAA GAT GGC AAA ACA CTA GTA TCA AAA       336
Thr Thr Leu Glu Val Phe Lys Glu Asp Gly Lys Thr Leu Val Ser Lys
            100                 105                 110

AAA GTA ACT TCC AAA GAC AAG TCA TCA ACA GAA GAA AAA TTC AAT GAA       384
Lys Val Thr Ser Lys Asp Lys Ser Ser Thr Glu Glu Lys Phe Asn Glu
        115                 120                 125

AAA GGT GAA GTA TCT GAA AAA ATA ATA ACA AGA GCA GAC GGA ACC AGA       432
Lys Gly Glu Val Ser Glu Lys Ile Ile Thr Arg Ala Asp Gly Thr Arg
    130                 135                 140

CTT GAA TAC ACA GGA ATT AAA AGC GAT GGA TCT GGA AAA GCT AAA GAG       480
Leu Glu Tyr Thr Gly Ile Lys Ser Asp Gly Ser Gly Lys Ala Lys Glu
145                 150                 155                 160

GTT TTA AAA GGC TAT GTT CTT GAA GGA ACT CTA ACT GCT GAA AAA ACA       528
Val Leu Lys Gly Tyr Val Leu Glu Gly Thr Leu Thr Ala Glu Lys Thr
                165                 170                 175

ACA TTG GTG GTT AAA GAA GGA ACT GTT ACT TTA AGC AAA AAT ATT TCA       576
Thr Leu Val Val Lys Glu Gly Thr Val Thr Leu Ser Lys Asn Ile Ser
            180                 185                 190

AAA TCT GGG GAA GTT TCA GTT GAA CTT AAT GAC ACT GAC AGT AGT GCT       624
Lys Ser Gly Glu Val Ser Val Glu Leu Asn Asp Thr Asp Ser Ser Ala
        195                 200                 205
```

Figure 7 (1 of 2)

```
GCT ACT AAA AAA ACT GCA GCT TGG AAT TCA GGC ACT TCA ACT TTA ACA      672
Ala Thr Lys Lys Thr Ala Ala Trp Asn Ser Gly Thr Ser Thr Leu Thr
    210                 215                 220

ATT ACT GTA AAC AGT AAA AAA ACT AAA GAC CTT GTG TTT ACA AAA GAA      720
Ile Thr Val Asn Ser Lys Lys Thr Lys Asp Leu Val Phe Thr Lys Glu
225                 230                 235                 240

AAC ACA ATT ACA GTA CAA CAA TAC GAC TCA AAT GGC ACC AAA TTA GAG      768
Asn Thr Ile Thr Val Gln Gln Tyr Asp Ser Asn Gly Thr Lys Leu Glu
                245                 250                 255

GGG TCA GCA GTT GAA ATT ACA AAA CTT GAT GAA ATT AAA AAC GCT TTA      816
Gly Ser Ala Val Glu Ile Thr Lys Leu Asp Glu Ile Lys Asn Ala Leu
            260                 265                 270

AAA TA                                                              822
Lys
```

Figure 7 (2 of 2)

OSPA K48

```
ATG AAA AAA TAT TTA TTG GGA ATA GGT CTA ATA TTA GCC TTA ATA GCA
 TAC TTT TTT ATA AAT AAC CCT TAT CCA GAT TAT AAT CGG AAT TAT CGT
 Met Lys Lys Tyr Leu Leu Gly Ile Gly Leu Ile Leu Ala Leu Ile Ala>

 50          60          70          80          90
 *     *     *     *     *     *     *     *     *     *
 TGT AAG CAA AAT GTT AGC AGC CTT GAT GAA AAA AAT AGC GTT TCA GTA
 ACA TTC GTT TTA CAA TCG TCG GAA CTA CTT TTT TTA TCG CAA AGT CAT
 Cys Lys Gln Asn Val Ser Ser Leu Asp Glu Lys Asn Ser Val Ser Val>

   100         110         120         130         140
    *     *     *     *     *     *     *     *     *
 GAT TTA CCT GGT GGA ATG ACA GTT CTT GTA AGT AAA GAA AAA GAC AAA
 CTA AAT GGA CCA CCT TAC TGT CAA GAA CAT TCA TTT CTT TTT CTG TTT
 Asp Leu Pro Gly Gly Met Thr Val Leu Val Ser Lys Glu Lys Asp Lys>

    150         160         170         180         190
 *     *     *     *     *     *     *     *     *     *
 GAC GGT AAA TAC AGT CTA GAG GCA ACA GTA GAC AAG CTT GAG CTT AAA
 CTG CCA TTT ATG TCA GAT CTC CGT TGT CAT CTG TTC GAA CTC GAA TTT
 Asp Gly Lys Tyr Ser Leu Glu Ala Thr Val Asp Lys Leu Glu Leu Lys>

       200         210         220         230         240
  *     *     *     *     *     *     *     *     *     *
 GGA ACT TCT GAT AAA AAC AAC GGT TCT GGA ACA CTT GAA GGT GAA AAA
 CCT TGA AGA CTA TTT TTG TTG CCA AGA CCT TGT GAA CTT CCA CTT TTT
 Gly Thr Ser Asp Lys Asn Asn Gly Ser Gly Thr Leu Glu Gly Glu Lys>

          250         260         270         280
    *     *     *     *     *     *     *     *     *
 ACT GAC AAA AGT AAA GTA AAA TTA ACA ATT GCT GAT GAC CTA AGT CAA
 TGA CTG TTT TCA TTT CAT TTT AAT TGT TAA CGA CTA CTG GAT TCA GTT
 Thr Asp Lys Ser Lys Val Lys Leu Thr Ile Ala Asp Asp Leu Ser Gln>

290         300         310         320         330
 *     *     *     *     *     *     *     *     *     *
 ACT AAA TTT GAA ATT TTC AAA GAA GAT GCC AAA ACA TTA GTA TCA AAA
 TGA TTT AAA CTT TAA AAG TTT CTT CTA CGG TTT TGT AAT CAT AGT TTT
 Thr Lys Phe Glu Ile Phe Lys Glu Asp Ala Lys Thr Leu Val Ser Lys>

   340         350         360         370         380
    *     *     *     *     *     *     *     *     *
 AAA GTA ACC CTT AAA GAC AAG TCA TCA ACA GAA GAA AAA TTC AAC GAA
 TTT CAT TGG GAA TTT CTG TTC AGT AGT TGT CTT CTT TTT AAG TTG CTT
 Lys Val Thr Leu Lys Asp Lys Ser Ser Thr Glu Glu Lys Phe Asn Glu>
```

FIGURE 8 (1 of 3)

OSP A K48

```
    390         400         410         420         430
*   *   *   *   *   *   *   *   *   *   *   *   *   *
AAG GGT GAA ACA TCT GAA AAA ACA ATA GTA AGA GCA AAT GGA ACC AGA
TTC CCA CTT TGT AGA CTT TTT TGT TAT CAT TCT CGT TTA CCT TGG TCT
Lys Gly Glu Thr Ser Glu Lys Thr Ile Val Arg Ala Asn Gly Thr Arg>

      440         450         460         470         480
  *   *   *   *   *   *   *   *   *   *   *   *   *   *
CTT GAA TAC ACA GAC ATA AAA AGC GAT GGA TCC GGA AAA GCT AAA GAA
GAA CTT ATG TGT CTG TAT TTT TCG CTA CCT AGG CCT TTT CGA TTT CTT
Leu Glu Tyr Thr Asp Ile Lys Ser Asp Gly Ser Gly Lys Ala Lys Glu>

          490         500         510         520
    *   *   *   *   *   *   *   *   *   *   *   *
GTT TTA AAA GAC TTT ACT CTT GAA GGA ACT CTA GCT GCT GAC GGC AAA
CAA AAT TTT CTG AAA TGA GAA CTT CCT TGA GAT CGA CGA CTG CCG TTT
Val Leu Lys Asp Phe Thr Leu Glu Gly Thr Leu Ala Ala Asp Gly Lys>

530         540         550         560         570
  *   *   *   *   *   *   *   *   *   *   *   *   *   *
ACA ACA TTG AAA GTT ACA GAA GGC ACT GTT GTT TTA AGC AAG AAC ATT
TGT TGT AAC TTT CAA TGT CTT CCG TGA CAA CAA AAT TCG TTC TTG TAA
Thr Thr Leu Lys Val Thr Glu Gly Thr Val Val Leu Ser Lys Asn Ile>

    580         590         600         610         620
  *   *   *   *   *   *   *   *   *   *   *   *   *   *
TTA AAA TCC GGA GAA ATA ACA GTT GCA CTT GAT GAC TCT GAC ACT ACT
AAT TTT AGG CCT CTT TAT TGT CAA CGT GAA CTA CTG AGA CTG TGA TGA
Leu Lys Ser Gly Glu Ile Thr Val Ala Leu Asp Asp Ser Asp Thr Thr>

    630         640         650         660         670
*   *   *   *   *   *   *   *   *   *   *   *   *   *
CAG GCT ACT AAA AAA ACT GGA AAA TGG GAT TCA AAA ACT TCC ACT TTA
GTC CGA TGA TTT TTT TGA CCT TTT ACC CTA AGT TTT TGA AGG TGA AAT
Gln Ala Thr Lys Lys Thr Gly Lys Trp Asp Ser Lys Thr Ser Thr Leu>

      680         690         700         710         720
  *   *   *   *   *   *   *   *   *   *   *   *   *   *
ACA ATT AGT GTG AAT AGC CAA AAA ACC AAA AAC CTT GTA TTC ACA AAA
TGT TAA TCA CAC TTA TCG GTT TTT TGG TTT TTG GAA CAT AAG TGT TTT
Thr Ile Ser Val Asn Ser Gln Lys Thr Lys Asn Leu Val Phe Thr Lys>

          730         740         750         760
    *   *   *   *   *   *   *   *   *   *   *   *   *
GAA GAC ACA ATA ACA GTA CAA AAA TAC GAC TCA GCA GGC ACC AAT CTA
CTT CTG TGT TAT TGT CAT GTT TTT ATG CTG AGT CGT CCG TGG TTA GAT
Glu Asp Thr Ile Thr Val Gln Lys Tyr Asp Ser Ala Gly Thr Asn Leu>
```

FIGURE 8 (2 of 3)

Osp A K-48

```
          770         780         790         800         810
           *     *     *     *     *     *     *     *     *     *
GAA GGC AAA GCA GTC GAA ATT ACA ACA CTT AAA GAA CTT AAA AAC GCT
CTT CCG TTT CGT CAG CTT TAA TGT TGT GAA TTT CTT GAA TTT TTG CGA
Glu Gly Lys Ala Val Glu Ile Thr Thr Leu Lys Glu Leu Lys Asn Ala>
```

OSP A K48

```
          820
           *     *
TTA AAA TAA
AAT TTT ATT
Leu Lys ***>
```

FIGURE 8 (3 of 3)

OSP A PGAU

```
          10          20          30          40
     *    *    *    *    *    *    *    *    *    *
ATG AAA AAA TAT TTA TTG GGA ATA GGT CTA ATA TTA GCC TTA ATA GCA
TAC TTT TTT ATA AAT AAC CCT TAT CCA GAT TAT AAT CGG AAT TAT CGT
Met Lys Lys Tyr Leu Leu Gly Ile Gly Leu Ile Leu Ala Leu Ile Ala>

50          60          70          80          90
 *    *    *    *    *    *    *    *    *    *
TGC AAG CAA AAT GTT AGC AGC CTT GAT GAA AAA AAC AGC GCT TCA GTA
ACG TTC GTT TTA CAA TCG TCG GAA CTA CTT TTT TTG TCG CGA AGT CAT
Cys Lys Gln Asn Val Ser Ser Leu Asp Glu Lys Asn Ser Ala Ser Val>

  100         110         120         130         140
   *    *    *    *    *    *    *    *    *    *
GAT TTG CCT GGT GAG ATG AAA GTT CTT GTA AGT AAA GAA AAA GAC AAA
CTA AAC GGA CCA CTC TAC TTT CAA GAA CAT TCA TTT CTT TTT CTG TTT
Asp Leu Pro Gly Glu Met Lys Val Leu Val Ser Lys Glu Lys Asp Lys>

    150         160         170         180         190
 *    *    *    *    *    *    *    *    *    *
GAC GGT AAG TAC AGT CTA AAG GCA ACA GTA GAC AAG ATT GAG CTA AAA
CTG CCA TTC ATG TCA GAT TTC CGT TGT CAT CTG TTC TAA CTC GAT TTT
Asp Gly Lys Tyr Ser Leu Lys Ala Thr Val Asp Lys Ile Glu Leu Lys>

       200         210         220         230         240
   *    *    *    *    *    *    *    *    *    *
GGA ACT TCT GAT AAA GAC AAT GGT TCT GGA GTG CTT GAA GGT ACA AAA
CCT TGA AGA CTA TTT CTG TTA CCA AGA CCT CAC GAA CTT CCA TGT TTT
Gly Thr Ser Asp Lys Asp Asn Gly Ser Gly Val Leu Glu Gly Thr Lys>

          250         260         270         280
     *    *    *    *    *    *    *    *    *    *
GAT GAC AAA AGT AAA GCA AAA TTA ACA ATT GCT GAC GAT CTA AGT AAA
CTA CTG TTT TCA TTT CGT TTT AAT TGT TAA CGA CTG CTA GAT TCA TTT
Asp Asp Lys Ser Lys Ala Lys Leu Thr Ile Ala Asp Asp Leu Ser Lys>

290         300         310         320         330
 *    *    *    *    *    *    *    *    *    *
ACC ACA TTC GAA CTT TTA AAA GAA GAT GGC AAA ACA TTA GTG TCA AGA
TGG TGT AAG CTT GAA AAT TTT CTT CTA CCG TTT TGT AAT CAC AGT TCT
Thr Thr Phe Glu Leu Leu Lys Glu Asp Gly Lys Thr Leu Val Ser Arg>

  340         350         360         370         380
   *    *    *    *    *    *    *    *    *    *
AAA GTA AGT TCT AGA GAC AAA ACA TCA ACA GAT GAA ATG TTC AAT GAA
TTT CAT TCA AGA TCT CTG TTT TGT AGT TGT CTA CTT TAC AAG TTA CTT
Lys Val Ser Ser Arg Asp Lys Thr Ser Thr Asp Glu Met Phe Asn Glu>
```

FIGURE 9 (1 of 3)

```
    390          400          410          420          430
    .    .       .    .       .    .       .    .       .    .
AAA GGT GAA TTG TCT GCA AAA ACC ATG ACA AGA GAA AAT GGA ACC AAA
TTT CCA CTT AAC AGA CGT TTT TGG TAC TGT TCT CTT TTA CCT TGG TTT
Lys Gly Glu Leu Ser Ala Lys Thr Met Thr Arg Glu Asn Gly Thr Lys>

      440          450          460          470          480
   .    .       .    .       .    .       .    .       .    .
CTT GAA TAT ACA GAA ATG AAA AGC GAT GGA ACC GGA AAA GCT AAA GAA
GAA CTT ATA TGT CTT TAC TTT TCG CTA CCT TGG CCT TTT CGA TTT CTT
Leu Glu Tyr Thr Glu Met Lys Ser Asp Gly Thr Gly Lys Ala Lys Glu>

          490          500          510          520
      .    .       .    .       .    .       .    .       .
GTT TTA AAA AAG TTT ACT CTT GAA GGA AAA GTA GCT AAT GAT AAA GTA
CAA AAT TTT TTC AAA TGA GAA CTT CCT TTT CAT CGA TTA CTA TTT CAT
Val Leu Lys Lys Phe Thr Leu Glu Gly Lys Val Ala Asn Asp Lys Val>

530          540          550          560          570
  .    .       .    .       .    .       .    .       .    .
ACA TTG GAA GTA AAA GAA GGA ACC GTT ACT TTA AGT AAG GAA ATT GCA
TGT AAC CTT CAT TTT CTT CCT TGG CAA TGA AAT TCA TTC CTT TAA CGT
Thr Leu Glu Val Lys Glu Gly Thr Val Thr Leu Ser Lys Glu Ile Ala>

  580          590          600          610          620
    .    .       .    .       .    .       .    .       .
AAA TCT GGA GAA GTA ACA GTT GCT CTT AAT GAC ACT AAC ACT ACT CAG
TTT AGA CCT CTT CAT TGT CAA CGA GAA TTA CTG TGA TTG TGA TGA GTC
Lys Ser Gly Glu Val Thr Val Ala Leu Asn Asp Thr Asn Thr Thr Gln>

    630          640          650          660          670
    .    .       .    .       .    .       .    .       .    .
GCT ACT AAA AAA ACT GGC GCA TGG GAT TCA AAA ACT TCT ACT TTA ACA
CGA TGA TTT TTT TGA CCG CGT ACC CTA AGT TTT TGA AGA TGA AAT TGT
Ala Thr Lys Lys Thr Gly Ala Trp Asp Ser Lys Thr Ser Thr Leu Thr>

      680          690          700          710          720
   .    .       .    .       .    .       .    .       .    .
ATT AGT GTT AAC AGC AAA AAA ACT ACA CAA CTT GTG TTT ACT AAA CAA
TAA TCA CAA TTG TCG TTT TTT TGA TGT GTT GAA CAC AAA TGA TTT GTT
Ile Ser Val Asn Ser Lys Lys Thr Thr Gln Leu Val Phe Thr Lys Gln>

          730          740          750          760
      .    .       .    .       .    .       .    .       .
TAC ACA ATA ACT GTA AAA CAA TAC GAC TCC GCA GGT ACC AAT TTA GAA
ATG TGT TAT TGA CAT TTT GTT ATG CTG AGG CGT CCA TGG TTA AAT CTT
Tyr Thr Ile Thr Val Lys Gln Tyr Asp Ser Ala Gly Thr Asn Leu Glu>
```

FIGURE 9 (2 of 3)

OSP A PGAU

```
          770        780        790        800        810
           *          *          *          *          *          *          *          *          *          *
GGC ACA GCA GTC GAA ATT AAA ACA CTT GAT GAA CTT AAA AAC GCT TTA
CCG TGT CGT CAG CTT TAA TTT TGT GAA CTA CTT GAA TTT TTG CGA AAT
Gly Thr Ala Val Glu Ile Lys Thr Leu Asp Glu Leu Lys Asn Ala Leu>

  820
   *
AAA TAA
TTT ATT
Lys ***>
```

FIGURE 9 (3 of 3)

```
ATG AAA AAA TAT TTA TTG GGA ATA GGT CTA ATA TTA GCT TTA ATA GCA        48
Met Lys Lys Tyr Leu Leu Gly Ile Gly Leu Ile Leu Ala Leu Ile Ala
  1               5                  10                  15

TGT AAG CAA AAT GTT AGC AGC CTT GAC GAG AAA AAC AGC GTT TCA GTA        96
Cys Lys Gln Asn Val Ser Ser Leu Asp Glu Lys Asn Ser Val Ser Val
            20                  25                  30

GAT TTG CCT GGT GAA ATG AAA GTT CTT GTA AGC AAA GAA AAA GAC AAA       144
Asp Leu Pro Gly Glu Met Lys Val Leu Val Ser Lys Glu Lys Asp Lys
        35                  40                  45

GAC GGC AAG TAC AGT CTA ATG GCA ACA GTA GAC AAG CTT GAG CTT AAA       192
Asp Gly Lys Tyr Ser Leu Met Ala Thr Val Asp Lys Leu Glu Leu Lys
    50                  55                  60
```

Figure 10 (1 of 2)

```
GGA ACA TCT GAT AAA AAC AAT GGA TCT GGG GTG CTT GAA GGC GTA AAA                240
Gly Thr Ser Asp Lys Asn Asn Gly Ser Gly Val Leu Glu Gly Val Lys
 65                  70                  75                  80

GCT GAC AAA AGC AAA GTA AAA TTA ACA GTT TCT GAC GAT CTA AGC ACA                288
Ala Asp Lys Ser Lys Val Lys Leu Thr Val Ser Asp Asp Leu Ser Thr
                 85                  90                  95

ACC ACA CTT GAA GTT TTA AAA GAA GAT GGC AAA ACA TTA GTG TCA AAA                336
Thr Thr Leu Glu Val Leu Lys Glu Asp Gly Lys Thr Leu Val Ser Lys
            100                 105                 110

AAA AGA ACT TCT AAA GAT AAG TCA TCA ACA GAA GAA AAG TTC AAT GAA                384
Lys Arg Thr Ser Lys Asp Lys Ser Ser Thr Glu Glu Lys Phe Asn Glu
        115                 120                 125

AAA GGC GAA TTA GTT GAA AAA ATA ATG GCA AGA GCA AAC GGA ACC ATA                432
Lys Gly Glu Leu Val Glu Lys Ile Met Ala Arg Ala Asn Gly Thr Ile
    130                 135                 140

CTT GAA TAC ACA GGA ATT AAA AGC GAT GGA TCC GGA AAA GCT AAA GAA                480
Leu Glu Tyr Thr Gly Ile Lys Ser Asp Gly Ser Gly Lys Ala Lys Glu
145                 150                 155                 160

ACT TTA AAA GAA TAT GTT CTT GAA GGA ACT CTA ACT GCT GAA AAA GCA                528
Thr Leu Lys Glu Tyr Val Leu Glu Gly Thr Leu Thr Ala Glu Lys Ala
                165                 170                 175

ACA TTG GTG GTT AAA GAA GGA ACT GTT ACT TTA AGT AAG CAC ATT TCA                576
Thr Leu Val Val Lys Glu Gly Thr Val Thr Leu Ser Lys His Ile Ser
            180                 185                 190

AAA TCT GGA GAA GTA ACA GCT GAA CTT AAT GAC ACT GAC AGT ACT CAA                624
Lys Ser Gly Glu Val Thr Ala Glu Leu Asn Asp Thr Asp Ser Thr Gln
        195                 200                 205

GCT ACT AAA AAA ACT GGG AAA TGG GAT GCA GGC ACT TCA ACT TTA ACA                672
Ala Thr Lys Lys Thr Gly Lys Trp Asp Ala Gly Thr Ser Thr Leu Thr
    210                 215                 220

ATT ACT GTA AAC AAC AAA AAA ACT AAA GCC CTT GTA TTT ACA AAA CAA                720
Ile Thr Val Asn Asn Lys Lys Thr Lys Ala Leu Val Phe Thr Lys Gln
225                 230                 235                 240

GAC ACA ATT ACA TCA CAA AAA TAC GAC TCA GCA GGA ACC AAC TTG GAA                768
Asp Thr Ile Thr Ser Gln Lys Tyr Asp Ser Ala Gly Thr Asn Leu Glu
                245                 250                 255

GGC ACA GCA GTC GAA ATT AAA ACA CTT GAT GAA CTT AAA AAC GCT TTA                816
Gly Thr Ala Val Glu Ile Lys Thr Leu Asp Glu Leu Lys Asn Ala Leu
            260                 265                 270

AGA                                                                           819
Arg
```

Figure 10 (2 of 2)

OSP B B-31
Sequence Range: 1 to 891

```
          10          20          30          40
     .    .    .    .    .    .    .    .    .    .
ATG AGA TTA TTA ATA GGA TTT GCT TTA GCG TTA GCT TTA ATA GGA TGT
TAC TCT AAT AAT TAT CCT AAA CGA AAT CGC AAT CGA AAT TAT CCT ACA
Met Arg Leu Leu Ile Gly Phe Ala Leu Ala Leu Ala Leu Ile Gly Cys>

50          60          70          80          90
.    .    .    .    .    .    .    .    .    .
GCA CAA AAA GGT GCT GAG TCA ATT GGT TCT CAA AAA GAA AAT GAT CTA
CGT GTT TTT CCA CGA CTC AGT TAA CCA AGA GTT TTT CTT TTA CTA GAT
Ala Gln Lys Gly Ala Glu Ser Ile Gly Ser Gln Lys Glu Asn Asp Leu>

  100         110         120         130         140
     .    .    .    .    .    .    .    .    .    .
AAC CTT GAA GAC TCT AGT AAA AAA TCA CAT CAA AAC GCT AAA CAA GAC
TTG GAA CTT CTG AGA TCA TTT TTT AGT GTA GTT TTG CGA TTT GTT CTG
Asn Leu Glu Asp Ser Ser Lys Lys Ser His Gln Asn Ala Lys Gln Asp>

    150         160         170         180         190
.    .    .    .    .    .    .    .    .    .
CTT CCT GCG GTG ACA GAA GAC TCA GTG TCT TTG TTT AAT GGT AAT AAA
GAA GGA CGC CAC TGT CTT CTG AGT CAC AGA AAC AAA TTA CCA TTA TTT
Leu Pro Ala Val Thr Glu Asp Ser Val Ser Leu Phe Asn Gly Asn Lys>

       200         210         220         230         240
     .    .    .    .    .    .    .    .    .    .
ATT TTT GTA AGC AAA GAA AAA AAT AGC TCC GGC AAA TAT GAT TTA AGA
TAA AAA CAT TCG TTT CTT TTT TTA TCG AGG CCG TTT ATA CTA AAT TCT
Ile Phe Val Ser Lys Glu Lys Asn Ser Ser Gly Lys Tyr Asp Leu Arg>

          250         260         270         280
     .    .    .    .    .    .    .    .    .    .
GCA ACA ATT GAT CAG GTT GAA CTT AAA GGA ACT TCC GAT AAA AAC AAT
CGT TGT TAA CTA GTC CAA CTT GAA TTT CCT TGA AGG CTA TTT TTG TTA
Ala Thr Ile Asp Gln Val Glu Leu Lys Gly Thr Ser Asp Lys Asn Asn>

290         300         310         320         330
.    .    .    .    .    .    .    .    .    .
GGT TCT GGA ACC CTT GAA GGT TCA AAG CCT GAC AAG AGT AAA GTA AAA
CCA AGA CCT TGG GAA CTT CCA AGT TTC GGA CTG TTC TCA TTT CAT TTT
Gly Ser Gly Thr Leu Glu Gly Ser Lys Pro Asp Lys Ser Lys Val Lys>

  340         350         360         370         380
     .    .    .    .    .    .    .    .    .    .
TTA ACA GTT TCT GCT GAT TTA AAC ACA GTA ACC TTA GAA GCA TTT GAT
AAT TGT CAA AGA CGA CTA AAT TTG TGT CAT TGG AAT CTT CGT AAA CTA
Leu Thr Val Ser Ala Asp Leu Asn Thr Val Thr Leu Glu Ala Phe Asp>
    390         400         410         420         430
```

FIGURE 11 (1 of 3)

```
                                                                  .
GCC AGC AAC CAA AAA ATT TCA AGT AAA GTT ACT AAA AAA CAG GGG TCA
CGG TCG TTG GTT TTT TAA AGT TCA TTT CAA TGA TTT TTT GTC CCC AGT
Ala Ser Asn Gln Lys Ile Ser Ser Lys Val Thr Lys Lys Gln Gly Ser>

      440         450         460         470         480
ATA ACA GAG GAA ACT CTC AAA GCT AAT AAA TTA GAC TCA AAG AAA TTA
TAT TGT CTC CTT TGA GAG TTT CGA TTA TTT AAT CTG AGT TTC TTT AAT
Ile Thr Glu Glu Thr Leu Lys Ala Asn Lys Leu Asp Ser Lys Lys Leu>

          490         500         510         520
ACA AGA TCA AAC GGA ACT ACA CTT GAA TAC TCA CAA ATA ACA GAT GCT
TGT TCT AGT TTG CCT TGA TGT GAA CTT ATG AGT GTT TAT TGT CTA CGA
Thr Arg Ser Asn Gly Thr Thr Leu Glu Tyr Ser Gln Ile Thr Asp Ala>

530         540         550         560         570
GAC AAT GCT ACA AAA GCA GTA GAA ACT CTA AAA AAT AGC ATT AAG CTT
CTG TTA CGA TGT TTT CGT CAT CTT TGA GAT TTT TTA TCG TAA TTC GAA
Asp Asn Ala Thr Lys Ala Val Glu Thr Leu Lys Asn Ser Ile Lys Leu>

  580         590         600         610         620
GAA GGA AGT CTT GTA GTC GGA AAA ACA ACA GTG GAA ATT AAA GAA GGT
CTT CCT TCA GAA CAT CAG CCT TTT TGT TGT CAC CTT TAA TTT CTT CCA
Glu Gly Ser Leu Val Val Gly Lys Thr Thr Val Glu Ile Lys Glu Gly>

    630         640         650         660         670
ACT GTT ACT CTA AAA AGA GAA ATT GAA AAA GAT GGA AAA GTA AAA GTC
TGA CAA TGA GAT TTT TCT CTT TAA CTT TTT CTA CCT TTT CAT TTT CAG
Thr Val Thr Leu Lys Arg Glu Ile Glu Lys Asp Gly Lys Val Lys Val>

      680         690         700         710         720
TTT TTG AAT GAC ACT GCA GGT TCT AAC AAA AAA ACA GGT AAA TGG GAA
AAA AAC TTA CTG TGA CGT CCA AGA TTG TTT TTT TGT CCA TTT ACC CTT
Phe Leu Asn Asp Thr Ala Gly Ser Asn Lys Lys Thr Gly Lys Trp Glu>

          730         740         750         760
GAC AGT ACT AGC ACT TTA ACA ATT AGT GCT GAC AGC AAA AAA ACT AAA
CTG TCA TGA TCG TGA AAT TGT TAA TCA CGA CTG TCG TTT TTT TGA TTT
Asp Ser Thr Ser Thr Leu Thr Ile Ser Ala Asp Ser Lys Lys Thr Lys>

770         780         790         800         810
GAT TTG GTG TTC TTA ACA GAT GGT ACA ATT ACA GTA CAA CAA TAC AAC
CTA AAC CAC AAG AAT TGT CTA CCA TGT TAA TGT CAT GTT GTT ATG TTG
Asp Leu Val Phe Leu Thr Asp Gly Thr Ile Thr Val Gln Gln Tyr Asn>
```

FIGURE 11 (2 of 3)

```
  820         830         840         850         860
   *     *     *     *     *     *     *     *     *
ACA GCT GGA ACC AGC CTA GAA GGA TCA GCA AGT GAA ATT AAA AAT CTT
TGT CGA CCT TGG TCG GAT CTT CCT AGT CGT TCA CTT TAA TTT TTA GAA
Thr Ala Gly Thr Ser Leu Glu Gly Ser Ala Ser Glu Ile Lys Asn Leu>

    870         880         890
*    *     *     *     *     *
TCA GAG CTT AAA AAC GCT TTA AAA TAA
AGT CTC GAA TTT TTG CGA AAT TTT ATT
Ser Glu Leu Lys Asn Ala Leu Lys ***>
```

FIGURE 11 (3 of 3)

OspC-B31
Sequence Range: 1 to 633

```
                10          20          30          40
                *           *           *           *           *
ATG AAA AAG AAT ACA TTA AGT GCG ATA TTA ATG ACT TTA TTT TTA TTT
TAC TTT TTC TTA TGT AAT TCA CGC TAT AAT TAC TGA AAT AAA AAT AAA
Met Lys Lys Asn Thr Leu Ser Ala Ile Leu Met Thr Leu Phe Leu Phe>

50          60          70          80          90
*           *           *           *           *           *
ATA TCT TGT AAT AAT TCA GGG AAA GAT GGG AAT ACA TCT GCA AAT TCT
TAT AGA ACA TTA TTA AGT CCC TTT CTA CCC TTA TGT AGA CGT TTA AGA
Ile Ser Cys Asn Asn Ser Gly Lys Asp Gly Asn Thr Ser Ala Asn Ser>

100         110         120         130         140
*           *           *           *           *           *
GCT GAT GAG TCT GTT AAA GGG CCT AAT CTT ACA GAA ATA AGT AAA AAA
CGA CTA CTC AGA CAA TTT CCC GGA TTA GAA TGT CTT TAT TCA TTT TTT
Ala Asp Glu Ser Val Lys Gly Pro Asn Leu Thr Glu Ile Ser Lys Lys>

150         160         170         180         190
*           *           *           *           *           *
ATT ACG GAT TCT AAT GCG GTT TTA CTT GCT GTG AAA GAG GTT GAA GCG
TAA TGC CTA AGA TTA CGC CAA AAT GAA CGA CAC TTT CTC CAA CTT CGC
Ile Thr Asp Ser Asn Ala Val Leu Leu Ala Val Lys Glu Val Glu Ala>

200         210         220         230         240
*           *           *           *           *           *
TTG CTG TCA TCT ATA GAT GAA ATT GCT GCT AAA GCT ATT GGT AAA AAA
AAC GAC AGT AGA TAT CTA CTT TAA CGA CGA TTT CGA TAA CCA TTT TTT
Leu Leu Ser Ser Ile Asp Glu Ile Ala Ala Lys Ala Ile Gly Lys Lys>

250         260         270         280
*           *           *           *           *           *
ATA CAC CAA AAT AAT GGT TTG GAT ACC GAA TAT AAT CAC AAT GGA TCA
TAT GTG GTT TTA TTA CCA AAC CTA TGG CTT ATA TTA GTG TTA CCT AGT
Ile His Gln Asn Asn Gly Leu Asp Thr Glu Tyr Asn His Asn Gly Ser>

290         300         310         320         330
*           *           *           *           *           *
TTG TTA GCG GGA CGT TAT GCA ATA TCA ACC CTA ATA AAA CAA AAA TTA
AAC AAT CGC CCT GCA ATA CGT TAT AGT TGG GAT TAT TTT GTT TTT AAT
Leu Leu Ala Gly Arg Tyr Ala Ile Ser Thr Leu Ile Lys Gln Lys Leu>

340         350         360         370         380
*           *           *           *           *           *
GAT GGA TTG AAA AAT GAA GGA TTA AAG GAA AAA ATT GAT GCG GCT AAG
CTA CCT AAC TTT TTA CTT CCT AAT TTC CTT TTT TAA CTA CGC CGA TTC
Asp Gly Leu Lys Asn Glu Gly Leu Lys Glu Lys Ile Asp Ala Ala Lys>
```

FIGURE 12 (1 of 2)

OspC-B31

```
    390          400          410          420          430
*    *    *    *    *    *    *    *    *    *    *
AAA TGT TCT GAA ACA TTT ACT AAT AAA TTA AAA GAA AAA CAC ACA GAT
TTT ACA AGA CTT TGT AAA TGA TTA TTT AAT TTT CTT TTT GTG TGT CTA
Lys Cys Ser Glu Thr Phe Thr Asn Lys Leu Lys Glu Lys His Thr Asp>

      440          450          460          470          480
 *    *    *    *    *    *    *    *    *    *    *
CTT GGT AAA GAA GGT GTT ACT GAT GCT GAT GCA AAA GAA GCC ATT TTA
GAA CCA TTT CTT CCA CAA TGA CTA CGA CTA CGT TTT CTT CGG TAA AAT
Leu Gly Lys Glu Gly Val Thr Asp Ala Asp Ala Lys Glu Ala Ile Leu>

          490          500          510          520
    *    *    *    *    *    *    *    *    *
AAA ACA AAT GGT ACT AAA ACT AAA GGT GCT GAA GAA CTT GGA AAA TTA
TTT TGT TTA CCA TGA TTT TGA TTT CCA CGA CTT CTT GAA CCT TTT AAT
Lys Thr Asn Gly Thr Lys Thr Lys Gly Ala Glu Glu Leu Gly Lys Leu>

530          540          550          560          570
 *    *    *    *    *    *    *    *    *    *
TTT GAA TCA GTA GAG GTC TTG TCA AAA GCA GCT AAA GAG ATG CTT GCT
AAA CTT AGT CAT CTC CAG AAC AGT TTT CGT CGA TTT CTC TAC GAA CGA
Phe Glu Ser Val Glu Val Leu Ser Lys Ala Ala Lys Glu Met Leu Ala>

  580          590          600          610          620
    *    *    *    *    *    *    *    *    *
AAT TCA GTT AAA GAG CTT ACA AGC CCT GTT GTG GCA GAA AGT CCA AAA
TTA AGT CAA TTT CTC GAA TGT TCG GGA CAA CAC CGT CTT TCA GGT TTT
Asn Ser Val Lys Glu Leu Thr Ser Pro Val Val Ala Glu Ser Pro Lys>

    630
*    *
AAA CCT TAA
TTT GGA ATT
Lys Pro ***>
```

FIGURE 12 (2 of 2)

OspC-K48
Sequence Range: 1 to 630

22/133

```
          10          20          30          40
     *    *    *    *    *    *    *    *    *
ATG AAA AAG AAT ACA TTA AGT GCG ATA TTA ATG ACT TTA TTT TTA TTT
TAC TTT TTC TTA TGT AAT TCA CGC TAT AAT TAC TGA AAT AAA AAT AAA
Met Lys Lys Asn Thr Leu Ser Ala Ile Leu Met Thr Leu Phe Leu Phe>

50          60          70          80          90
 *    *    *    *    *    *    *    *    *    *
ATA TCT TGT AAT AAT TCA GGT GGG GAT ACC GCA TCT ACT AAT CCT GAT
TAT AGA ACA TTA TTA AGT CCA CCC CTA TGG CGT AGA TGA TTA GGA CTA
Ile Ser Cys Asn Asn Ser Gly Gly Asp Thr Ala Ser Thr Asn Pro Asp>

  100         110         120         130         140
   *    *    *    *    *    *    *    *    *
GAG TCT GCA AAA GGA CCT AAT CTT ACA GTA ATA AGC AAA AAA ATT ACA
CTC AGA CGT TTT CCT GGA TTA GAA TGT CAT TAT TCG TTT TTT TAA TGT
Glu Ser Ala Lys Gly Pro Asn Leu Thr Val Ile Ser Lys Lys Ile Thr>

    150         160         170         180         190
*    *    *    *    *    *    *    *    *    *
GAT TCT AAT GCA TTT GTA CTG GCT GTG AAA GAA GTT GAG GCT TTG ATC
CTA AGA TTA CGT AAA CAT GAC CGA CAC TTT CTT CAA CTC CGA AAC TAG
Asp Ser Asn Ala Phe Val Leu Ala Val Lys Glu Val Glu Ala Leu Ile>

      200         210         220         230         240
  *    *    *    *    *    *    *    *    *    *
TCA TCT ATA GAT GAA CTT GCT AAT AAA GCT ATT GGT AAA GTA ATA CAT
AGT AGA TAT CTA CTT GAA CGA TTA TTT CGA TAA CCA TTT CAT TAT GTA
Ser Ser Ile Asp Glu Leu Ala Asn Lys Ala Ile Gly Lys Val Ile His>

          250         260         270         280
     *    *    *    *    *    *    *    *    *
CAA AAT AAT GGT TTA AAT GCT AAT GCG GGT CAA AAC GGA TCA TTG TTA
GTT TTA TTA CCA AAT TTA CGA TTA CGC CCA GTT TTG CCT AGT AAC AAT
Gln Asn Asn Gly Leu Asn Ala Asn Ala Gly Gln Asn Gly Ser Leu Leu>

290         300         310         320         330
 *    *    *    *    *    *    *    *    *    *
GCA GGA GCC TAT GCA ATA TCA ACC CTA ATA ACA GAA AAA TTA AGT AAA
CGT CCT CGG ATA CGT TAT AGT TGG GAT TAT TGT CTT TTT AAT TCA TTT
Ala Gly Ala Tyr Ala Ile Ser Thr Leu Ile Thr Glu Lys Leu Ser Lys>

  340         350         360         370         380
   *    *    *    *    *    *    *    *    *
TTG AAA AAT TCA GAA GAG TTA AAT AAA AAA ATT GAA GAG GCT AAG AAC
AAC TTT TTA AGT CTT CTC AAT TTA TTT TTT TAA CTT CTC CGA TTC TTG
Leu Lys Asn Ser Glu Glu Leu Asn Lys Lys Ile Glu Glu Ala Lys Asn>
```

FIGURE 13 (1 of 2)

OspC-K48

```
    390           400           410           420           430
*     *     *     *     *     *     *     *     *     *     *
CAT TCT GAA GCA TTT ACT AAT AGA CTA AAA GGT TCT CAT GCA CAA CTT
GTA AGA CTT CGT AAA TGA TTA TCT GAT TTT CCA AGA GTA CGT GTT GAA
His Ser Glu Ala Phe Thr Asn Arg Leu Lys Gly Ser His Ala Gln Leu>

      440           450           460           470           480
  *     *     *     *     *     *     *     *     *     *     *
GGA GTT GCT GCT GCT ACT GAT GAT CAT GCA AAA GAA GCT ATT TTA AAG
CCT CAA CGA CGA CGA TGA CTA CTA GTA CGT TTT CTT CGA TAA AAT TTC
Gly Val Ala Ala Ala Thr Asp Asp His Ala Lys Glu Ala Ile Leu Lys>

         490           500           510           520
     *     *     *     *     *     *     *     *     *     *
TCA AAT CCT ACT AAA GAT AAG GGT GCT AAA GCA CTT AAA GAC TTA TCT
AGT TTA GGA TGA TTT CTA TTC CCA CGA TTT CGT GAA TTT CTG AAT AGA
Ser Asn Pro Thr Lys Asp Lys Gly Ala Lys Ala Leu Lys Asp Leu Ser>

530           540           550           560           570
  *     *     *     *     *     *     *     *     *     *
GAA TCA GTA GAA AGC TTG GCA AAA GCA GCG CAA GAA GCA TTA GCT AAT
CTT AGT CAT CTT TCG AAC CGT TTT CGT CGC GTT CTT CGT AAT CGA TTA
Glu Ser Val Glu Ser Leu Ala Lys Ala Ala Gln Glu Ala Leu Ala Asn>

  580           590           600           610           620
    *     *     *     *     *     *     *     *     *
TCA GTT AAA GAA CTT ACA AAT CCT GTT GTG GCA GAA AGT CCA AAA AAA
AGT CAA TTT CTT GAA TGT TTA GGA CAA CAC CGT CTT TCA GGT TTT TTT
Ser Val Lys Glu Leu Thr Asn Pro Val Val Ala Glu Ser Pro Lys Lys>

    630
*     *
CCT TAA
GGA ATT
Pro ***>
```

FIGURE 13 (2 of 2)

OspC-PKO
Sequence Range: 1 to 639

```
          10          20          30          40
     *    *    *    *    *    *    *    *    *
ATG AAA AAG AAT ACA TTA AGT GCG ATA TTA ATG ACT TTA TTT TTA TTT
TAC TTT TTC TTA TGT AAT TCA CGC TAT AAT TAC TGA AAT AAA AAT AAA
Met Lys Lys Asn Thr Leu Ser Ala Ile Leu Met Thr Leu Phe Leu Phe>

50          60          70          80          90
 *    *    *    *    *    *    *    *    *    *
ATA TCT TGT AGT AAT TCA GGG AAA GGT GGG GAT TCT GCA TCT ACT AAT
TAT AGA ACA TCA TTA AGT CCC TTT CCA CCC CTA AGA CGT AGA TGA TTA
Ile Ser Cys Ser Asn Ser Gly Lys Gly Gly Asp Ser Ala Ser Thr Asn>

  100         110         120         130         140
    *    *    *    *    *    *    *    *    *
CCT GCT GAC GAG TCT GCG AAA GGG CCT AAT CTT ACA GAA ATA AGC AAA
GGA CGA CTG CTC AGA CGC TTT CCC GGA TTA GAA TGT CTT TAT TCG TTT
Pro Ala Asp Glu Ser Ala Lys Gly Pro Asn Leu Thr Glu Ile Ser Lys>

    150         160         170         180         190
 *    *    *    *    *    *    *    *    *    *
AAA ATT ACA GAT TCT AAT GCA TTT GTA CTT GCT GTT AAA GAA GTT GAG
TTT TAA TGT CTA AGA TTA CGT AAA CAT GAA CGA CAA TTT CTT CAA CTC
Lys Ile Thr Asp Ser Asn Ala Phe Val Leu Ala Val Lys Glu Val Glu>

       200         210         220         230         240
  *    *    *    *    *    *    *    *    *    *
ACT TTG GTT TTA TCT ATA GAT GAA CTT GCT AAG AAA GCT ATT GGT CAA
TGA AAC CAA AAT AGA TAT CTA CTT GAA CGA TTC TTT CGA TAA CCA GTT
Thr Leu Val Leu Ser Ile Asp Glu Leu Ala Lys Lys Ala Ile Gly Gln>

          250         260         270         280
     *    *    *    *    *    *    *    *    *
AAA ATA GAC AAT AAT AAT GGT TTA GCT GCT TTA AAT AAT CAG AAT GGA
TTT TAT CTG TTA TTA TTA CCA AAT CGA CGA AAT TTA TTA GTC TTA CCT
Lys Ile Asp Asn Asn Asn Gly Leu Ala Ala Leu Asn Asn Gln Asn Gly>

290         300         310         320         330
 *    *    *    *    *    *    *    *    *    *
TCG TTG TTA GCA GGA GCC TAT GCA ATA TCA ACC CTA ATA ACA GAA AAA
AGC AAC AAT CGT CCT CGG ATA CGT TAT AGT TGG GAT TAT TGT CTT TTT
Ser Leu Leu Ala Gly Ala Tyr Ala Ile Ser Thr Leu Ile Thr Glu Lys>

  340         350         360         370         380
    *    *    *    *    *    *    *    *    *
TTG AGT AAA TTG AAA AAT TTA GAA GAA TTA AAG ACA GAA ATT GCA AAG
AAC TCA TTT AAC TTT TTA AAT CTT CTT AAT TTC TGT CTT TAA CGT TTC
Leu Ser Lys Leu Lys Asn Leu Glu Glu Leu Lys Thr Glu Ile Ala Lys>
```

FIGURE 14 (1 of 2)

OspC-PKO

```
    390         400         410         420         430
*   *   *   *   *   *   *   *   *   *   *   *   *   *
GCT AAG AAA TGT TCC GAA GAA TTT ACT AAT AAA CTA AAA AGT GGT CAT
CGA TTC TTT ACA AGG CTT CTT AAA TGA TTA TTT GAT TTT TCA CCA GTA
Ala Lys Lys Cys Ser Glu Glu Phe Thr Asn Lys Leu Lys Ser Gly His>

      440         450         460         470         480
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GCA GAT CTT GGC AAA CAG GAT GCT ACC GAT GAT CAT GCA AAA GCA GCT
CGT CTA GAA CCG TTT GTC CTA CGA TGG CTA CTA GTA CGT TTT CGT CGA
Ala Asp Leu Gly Lys Gln Asp Ala Thr Asp Asp His Ala Lys Ala Ala>

        490         500         510         520
*   *   *   *   *   *   *   *   *   *   *   *   *   *
ATT TTA AAA ACA CAT GCA ACT ACC GAT AAA GGT GCT AAA GAA TTT AAA
TAA AAT TTT TGT GTA CGT TGA TGG CTA TTT CCA CGA TTT CTT AAA TTT
Ile Leu Lys Thr His Ala Thr Thr Asp Lys Gly Ala Lys Glu Phe Lys>

530         540         550         560         570
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GAT TTA TTT GAA TCA GTA GAA GGT TTG TTA AAA GCA GCT CAA GTA GCA
CTA AAT AAA CTT AGT CAT CTT CCA AAC AAT TTT CGT CGA GTT CAT CGT
Asp Leu Phe Glu Ser Val Glu Gly Leu Leu Lys Ala Ala Gln Val Ala>

  580         590         600         610         620
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *
CTA ACT AAT TCA GTT AAA GAA CTT ACA AGT CCT GTT GTA GCA GAA AGT
GAT TGA TTA AGT CAA TTT CTT GAA TGT TCA GGA CAA CAT CGT CTT TCA
Leu Thr Asn Ser Val Lys Glu Leu Thr Ser Pro Val Val Ala Glu Ser>

    630
*   *   *
CCA AAA AAA CCT TAA
GGT TTT TTT GGA ATT
Pro Lys Lys Pro ***>
```

FIGURE 14 (2 of 2)

OspC-TRO
Sequence Range: 1 to 624

```
          10          20          30          40
     *    *    *    *    *    *    *    *    *
ATG AAA AAG AAT ACA TTA AGT GCG ATA TTA ATG ACT TTA TTT TTA TTT
TAC TTT TTC TTA TGT AAT TCA CGC TAT AAT TAC TGA AAT AAA AAT AAA
Met Lys Lys Asn Thr Leu Ser Ala Ile Leu Met Thr Leu Phe Leu Phe>

50          60          70          80          90
 *    *    *    *    *    *    *    *    *    *
ATA TCT TGT AAT AAT TCA GGT GGG GAT TCT GCA TCT ACT AAT CCT GAT
TAT AGA ACA TTA TTA AGT CCA CCC CTA AGA CGT AGA TGA TTA GGA CTA
Ile Ser Cys Asn Asn Ser Gly Gly Asp Ser Ala Ser Thr Asn Pro Asp>

  100         110         120         130         140
    *    *    *    *    *    *    *    *    *
GAG TCT GCA AAA GGA CCT AAT CTT ACC GTA ATA AGC AAA AAA ATT ACA
CTC AGA CGT TTT CCT GGA TTA GAA TGG CAT TAT TCG TTT TTT TAA TGT
Glu Ser Ala Lys Gly Pro Asn Leu Thr Val Ile Ser Lys Lys Ile Thr>

    150         160         170         180         190
 *    *    *    *    *    *    *    *    *    *
GAT TCT AAT GCA TTT TTA CTG GCT GTG AAA GAA GTT GAG GCT TTG CTT
CTA AGA TTA CGT AAA AAT GAC CGA CAC TTT CTT CAA CTC CGA AAC GAA
Asp Ser Asn Ala Phe Leu Leu Ala Val Lys Glu Val Glu Ala Leu Leu>

      200         210         220         230         240
   *    *    *    *    *    *    *    *    *    *
TCA TCT ATA GAT GAA CTT TCT AAA GCT ATT GGT AAA AAA ATA AAA AAT
AGT AGA TAT CTA CTT GAA AGA TTT CGA TAA CCA TTT TTT TAT TTT TTA
Ser Ser Ile Asp Glu Leu Ser Lys Ala Ile Gly Lys Lys Ile Lys Asn>

          250         260         270         280
     *    *    *    *    *    *    *    *    *
GAT GGT ACT TTA GAT AAC GAA GCA AAT CGA AAC GAA TCA TTG ATA GCA
CTA CCA TGA AAT CTA TTG CTT CGT TTA GCT TTG CTT AGT AAC TAT CGT
Asp Gly Thr Leu Asp Asn Glu Ala Asn Arg Asn Glu Ser Leu Ile Ala>

290         300         310         320         330
 *    *    *    *    *    *    *    *    *    *
GGA GCT TAT GAA ATA TCA AAA CTA ATA ACA CAA AAA TTA AGT GTA TTG
CCT CGA ATA CTT TAT AGT TTT GAT TAT TGT GTT TTT AAT TCA CAT AAC
Gly Ala Tyr Glu Ile Ser Lys Leu Ile Thr Gln Lys Leu Ser Val Leu>

  340         350         360         370         380
    *    *    *    *    *    *    *    *    *
AAT TCA GAA GAA TTA AAG AAA AAA ATT AAA GAG GCT AAG GAT TGT TCC
TTA AGT CTT CTT AAT TTC TTT TTT TAA TTT CTC CGA TTC CTA ACA AGG
Asn Ser Glu Glu Leu Lys Lys Lys Ile Lys Glu Ala Lys Asp Cys Ser>
```

FIGURE 15 (1 of 2)

OspC-TRO

```
    390         400         410         420         430
     *   *   *   *   *   *   *   *   *   *   *   *   *
GAA AAA TTT ACT ACT AAG CTA AAA GAT AGT CAT GCA GAG CTT GGT ATA
CTT TTT AAA TGA TGA TTC GAT TTT CTA TCA GTA CGT CTC GAA CCA TAT
Glu Lys Phe Thr Thr Lys Leu Lys Asp Ser His Ala Glu Leu Gly Ile>

      440         450         460         470         480
   *   *   *   *   *   *   *   *   *   *   *   *   *   *
CAA AGC GTT CAG GAT GAT AAT GCA AAA AAA GCT ATT TTA AAA ACA CAT
GTT TCG CAA GTC CTA CTA TTA CGT TTT TTT CGA TAA AAT TTT TGT GTA
Gln Ser Val Gln Asp Asp Asn Ala Lys Lys Ala Ile Leu Lys Thr His>

         490         500         510         520
   *   *   *   *   *   *   *   *   *   *   *   *   *
GGA ACT AAA GAC AAG GGT GCT AAA GAA CTT GAA GAG TTA TTT AAA TCA
CCT TGA TTT CTG TTC CCA CGA TTT CTT GAA CTT CTC AAT AAA TTT AGT
Gly Thr Lys Asp Lys Gly Ala Lys Glu Leu Glu Glu Leu Phe Lys Ser>

530         540         550         560         570
   *   *   *   *   *   *   *   *   *   *   *   *   *
CTA GAA AGC TTG TCA AAA GCA GCG CAA GCA GCA TTA ACT AAT TCA GTT
GAT CTT TCG AAC AGT TTT CGT CGC GTT CGT CGT AAT TGA TTA AGT CAA
Leu Glu Ser Leu Ser Lys Ala Ala Gln Ala Ala Leu Thr Asn Ser Val>

   580         590         600         610         620
   *   *   *   *   *   *   *   *   *   *   *   *   *
AAA GAG CTT ACA AAT CCT GTT GTG GCA GAA AGT CCA AAA AAA CCT TAA
TTT CTC GAA TGT TTA GGA CAA CAC CGT CTT TCA GGT TTT TTT GGA ATT
Lys Glu Leu Thr Asn Pro Val Val Ala Glu Ser Pro Lys Lys Pro ***>
```

FIGURE 15 (2 of 2)

P93
Sequence Range: 1 to 2102

```
                10          20          30          40
                *   *   *   *   *   *   *   *   *   *
ATG AAA AAA ATG TTA CTA ATC TTT AGT TTT TTT CTT ATT TTC TTG AAT
TAC TTT TTT TAC AAT GAT TAG AAA TCA AAA AAA GAA TAA AAG AAC TTA
Met Lys Lys Met Leu Leu Ile Phe Ser Phe Phe Leu Ile Phe Leu Asn>

50          60          70          80          90
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GGA TTT CCT GTT AGT GCA AGA GAA GTT GAT AGG GAA AAA TTA AAG GAC
CCT AAA GGA CAA TCA CGT TCT CTT CAA CTA TCC CTT TTT AAT TTC CTG
Gly Phe Pro Val Ser Ala Arg Glu Val Asp Arg Glu Lys Leu Lys Asp>

  100         110         120         130         140
  *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
TTT GTT AAT ATG GAT CTT GAG TTT GTA AAT TAT AAA GGC CCT TAT GAT
AAA CAA TTA TAC CTA GAA CTC AAA CAT TTA ATA TTT CCG GGA ATA CTA
Phe Val Asn Met Asp Leu Glu Phe Val Asn Tyr Lys Gly Pro Tyr Asp>

    150         160         170         180         190
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *
TCT ACA AAT ACA TAT GAA CAA ATA GTG GGT ATT GGG GAG TTT TTA GCA
AGA TGT TTA TGT ATA CTT GTT TAT CAC CCA TAA CCC CTC AAA AAT CGT
Ser Thr Asn Thr Tyr Glu Gln Ile Val Gly Ile Gly Glu Phe Leu Ala>

      200         210         220         230         240
  *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
AGA CCG TTG ACC AAT TCC AAT AGC AAC TCA AGT TAT TAT GGT AAA TAT
TCT GGC AAC TGG TTA AGG TTA TCG TTG AGT TCA ATA ATA CCA TTT ATA
Arg Pro Leu Thr Asn Ser Asn Ser Asn Ser Ser Tyr Tyr Gly Lys Tyr>

          250         260         270         280
      *   *   *   *   *   *   *   *   *   *   *   *   *   *
TTT ATT AAT AGA TTT ATT GAT GAT CAA GAT AAA AAA GCA AGC GTT GAT
AAA TAA TTA TCT AAA TAA CTA CTA GTT CTA TTT TTT CGT TCG CAA CTA
Phe Ile Asn Arg Phe Ile Asp Asp Gln Asp Lys Lys Ala Ser Val Asp>

290         300         310         320         330
*   *   *   *   *   *   *   *   *   *   *   *   *   *   *
GTT TTT TCT ATT GGT AGT AAG TCA GAG CTT GAC AGT ATA TTG AAT TTA
CAA AAA AGA TAA CCA TCA TTC AGT CTC GAA CTG TCA TAT AAC TTA AAT
Val Phe Ser Ile Gly Ser Lys Ser Glu Leu Asp Ser Ile Leu Asn Leu>

  340         350         360         370         380
  *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
AGA AGA ATT CTT ACA GGG TAT TTA ATA AAG TCT TTC GAT TAT GAC AGG
TCT TCT TAA GAA TGT CCC ATA AAT TAT TTC AGA AAG CTA ATA CTG TCC
Arg Arg Ile Leu Thr Gly Tyr Leu Ile Lys Ser Phe Asp Tyr Asp Arg>
```

FIGURE 16 (1 of 5)

```
    390         400         410         420         430
    .   .   .   .   .   .   .   .   .   .   .   .   .
TCT AGT GCA GAA TTA ATT GCT AAG GTT ATT ACA ATA TAT AAT GCT GTT
AGA TCA CGT CTT AAT TAA CGA TTC CAA TAA TGT TAT ATA TTA CGA CAA
Ser Ser Ala Glu Leu Ile Ala Lys Val Ile Thr Ile Tyr Asn Ala Val>

        440         450         460         470         480
    .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
TAT AGA GGA GAT TTG GAT TAT TAT AAA GGG TTT TAT ATT GAG GCT GCT
ATA TCT CCT CTA AAC CTA ATA ATA TTT CCC AAA ATA TAA CTC CGA CGA
Tyr Arg Gly Asp Leu Asp Tyr Tyr Lys Gly Phe Tyr Ile Glu Ala Ala>

          490         500         510         520
    .   .   .   .   .   .   .   .   .   .   .   .   .
TTA AAG TCT TTA AGT AAA GAA AAT GCA GGT CTT TCT AGG GTT TAT AGT
AAT TTC AGA AAT TCA TTT CTT TTA CGT CCA GAA AGA TCC CAA ATA TCA
Leu Lys Ser Leu Ser Lys Glu Asn Ala Gly Leu Ser Arg Val Tyr Ser>

530         540         550         560         570
    .   .   .   .   .   .   .   .   .   .   .   .   .
CAG TGG GCT GGA AAG ACA CAA ATA TTT ATT CCT CTT AAA AAG GAT ATT
GTC ACC CGA CCT TTC TGT GTT TAT AAA TAA GGA GAA TTT TTC CTA TAA
Gln Trp Ala Gly Lys Thr Gln Ile Phe Ile Pro Leu Lys Lys Asp Ile>

  580         590         600         610         620
    .   .   .   .   .   .   .   .   .   .   .   .   .
TTG TCT GGA AAT ATT GAG TCT GAC ATT GAT ATT GAC AGT TTA GTT ACA
AAC AGA CCT TTA TAA CTC AGA CTG TAA CTA TAA CTG TCA AAT CAA TGT
Leu Ser Gly Asn Ile Glu Ser Asp Ile Asp Ile Asp Ser Leu Val Thr>

    630         640         650         660         670
    .   .   .   .   .   .   .   .   .   .   .   .   .
GAT AAG GTG GTG GCA GCT CTT TTA AGT GAA AAT GAA GCA GGT GTT AAC
CTA TTC CAC CAC CGT CGA GAA AAT TCA CTT TTA CTT CGT CCA CAA TTG
Asp Lys Val Val Ala Ala Leu Leu Ser Glu Asn Glu Ala Gly Val Asn>

        680         690         700         710         720
    .   .   .   .   .   .   .   .   .   .   .   .   .   .   .
TTT GCA AGA GAT ATT ACA GAT ATT CAA GGC GAA ACT CAT AAG GCA GAT
AAA CGT TCT CTA TAA TGT CTA TAA GTT CCG CTT TGA GTA TTC CGT CTA
Phe Ala Arg Asp Ile Thr Asp Ile Gln Gly Glu Thr His Lys Ala Asp>

          730         740         750         760
    .   .   .   .   .   .   .   .   .   .   .   .   .
CAA GAT AAA ATT GAT ATT GAA TTA GAC AAT ATT CAT GAA AGT GAT TCC
GTT CTA TTT TAA CTA TAA CTT AAT CTG TTA TAA GTA CTT TCA CTA AGG
Gln Asp Lys Ile Asp Ile Glu Leu Asp Asn Ile His Glu Ser Asp Ser>

770         780         790         800         810
    .   .   .   .   .   .   .   .   .   .   .   .   .
AAT ATA ACA GAA ACT ATT GAA AAT TTA AGG GAT CAG CTT GAA AAA GCT
TTA TAT TGT CTT TGA TAA CTT TTA AAT TCC CTA GTC GAA CTT TTT CGA
Asn Ile Thr Glu Thr Ile Glu Asn Leu Arg Asp Gln Leu Glu Lys Ala>
```

FIGURE 16 (2 of 5)

```
820         830         840         850         860
ACA GAT GAA GAG CAT AAA AAA GAG ATT GAA AGT CAG GTT GAT GCT AAA
TGT CTA CTT CTC GTA TTT TTT CTC TAA CTT TCA GTC CAA CTA CGA TTT
Thr Asp Glu Glu His Lys Lys Glu Ile Glu Ser Gln Val Asp Ala Lys>

    870         880         890         900         910
AAG AAA CAA AAG GAA GAG CTA GAT AAA AAG GCA ATA AAT CTT GAT AAA
TTC TTT GTT TTC CTT CTC GAT CTA TTT TTC CGT TAT TTA GAA CTA TTT
Lys Lys Gln Lys Glu Glu Leu Asp Lys Lys Ala Ile Asn Leu Asp Lys>

       920         930         940         950         960
GCT CAG CAA AAA TTA GAT TCT GCT GAA GAT AAT TTA GAT GTT CAA AGA
CGA GTC GTT TTT AAT CTA AGA CGA CTT CTA TTA AAT CTA CAA GTT TCT
Ala Gln Gln Lys Leu Asp Ser Ala Glu Asp Asn Leu Asp Val Gln Arg>

          970         980         990        1000
AAT ACT GTT AGA GAG AAA ATT CAA GAG GAT ATT AAC GAA ATT AAC AAG
TTA TGA CAA TCT CTC TTT TAA GTT CTC CTA TAA TTG CTT TAA TTG TTC
Asn Thr Val Arg Glu Lys Ile Gln Glu Asp Ile Asn Glu Ile Asn Lys>

1010        1020        1030        1040        1050
GAA AAG AAT TTA CCA AAG CCT GGT GAT GTA AGT TCT CCT AAA GTT GAT
CTT TTC TTA AAT GGT TTC GGA CCA CTA CAT TCA AGA GGA TTT CAA CTA
Glu Lys Asn Leu Pro Lys Pro Gly Asp Val Ser Ser Pro Lys Val Asp>

 1060        1070        1080        1090        1100
AAG CAA CTA CAA ATA AAA GAG AGC CTG GAA GAT TTG CAG GAG CAG CTT
TTC GTT GAT GTT TAT TTT CTC TCG GAC CTT CTA AAC GTC CTC GTC GAA
Lys Gln Leu Gln Ile Lys Glu Ser Leu Glu Asp Leu Gln Glu Gln Leu>

   1110        1120        1130        1140        1150
AAA GAA ACT GGT GAT GAA AAT CAG AAA AGA GAA ATT GAA AAG CAA ATT
TTT CTT TGA CCA CTA CTT TTA GTC TTT TCT CTT TAA CTT TTC GTT TAA
Lys Glu Thr Gly Asp Glu Asn Gln Lys Arg Glu Ile Glu Lys Gln Ile>

       1160        1170        1180        1190        1200
GAA ATC AAA AAA AGT GAT GAA AAG CTT TTA AAA AGT AAA GAT GAT AAA
CTT TAG TTT TTT TCA CTA CTT TTC GAA AAT TTT TCA TTT CTA CTA TTT
Glu Ile Lys Lys Ser Asp Glu Lys Leu Leu Lys Ser Lys Asp Asp Lys>

          1210        1220        1230        1240
GCA AGT AAA GAT GGT AAA GCC TTG GAT CTT GAT CGA GAA TTA AAT TCT
CGT TCA TTT CTA CCA TTT CGG AAC CTA GAA CTA GCT CTT AAT TTA AGA
Ala Ser Lys Asp Gly Lys Ala Leu Asp Leu Asp Arg Glu Leu Asn Ser>
```

FIGURE 16 (3 of 5)

```
1250        1260        1270        1280        1290
AAA GCT TCT AGC AAA GAA AAA AGT AAA GCC AAG GAA GAA GAA ATA ACC
TTT CGA AGA TCG TTT CTT TTT TCA TTT CGG TTC CTT CTT CTT TAT TGG
Lys Ala Ser Ser Lys Glu Lys Ser Lys Ala Lys Glu Glu Glu Ile Thr>

1300        1310        1320        1330        1340
AAG GGT AAG TCA CAG AAA AGC TTA GGC GAT TTG AAT AAT GAT GAA AAT
TTC CCA TTC AGT GTC TTT TCG AAT CCG CTA AAC TTA TTA CTA CTT TTA
Lys Gly Lys Ser Gln Lys Ser Leu Gly Asp Leu Asn Asn Asp Glu Asn>

1350        1360        1370        1380        1390
CTT ATG ATG CCA GAA GAT CAA AAA TTA CCT GAG GTT AAA AAA TTA GAT
GAA TAC TAC GGT CTT CTA GTT TTT AAT GGA CTC CAA TTT TTT AAT CTA
Leu Met Met Pro Glu Asp Gln Lys Leu Pro Glu Val Lys Lys Leu Asp>

1400        1410        1420        1430        1440
AGC AAA AAA GAA TTT AAA CCT GTT TCT GAG GTT GAG AAA TTA GAT AAG
TCG TTT TTT CTT AAA TTT GGA CAA AGA CTC CAA CTC TTT AAT CTA TTC
Ser Lys Lys Glu Phe Lys Pro Val Ser Glu Val Glu Lys Leu Asp Lys>

1450        1460        1470        1480
ATT TTC AAG TCT AAT AAC AAT GTT GGA GAA TTA TCA CCG TTA GAT AAA
TAA AAG TTC AGA TTA TTG TTA CAA CCT CTT AAT AGT GGC AAT CTA TTT
Ile Phe Lys Ser Asn Asn Asn Val Gly Glu Leu Ser Pro Leu Asp Lys>

1490        1500        1510        1520        1530
TCT TCT TAT AAA GAC ATT GAT TCA AAA GAG GAG ACA GTT AAT AAA GAT
AGA AGA ATA TTT CTG TAA CTA AGT TTT CTC CTC TGT CAA TTA TTT CTA
Ser Ser Tyr Lys Asp Ile Asp Ser Lys Glu Glu Thr Val Asn Lys Asp>

1540        1550        1560        1570        1580
GTT AAT TTG CAA AAG ACT AAG CCT CAG GTT AAA GAC CAA GTT ACT TCT
CAA TTA AAC GTT TTC TGA TTC GGA GTC CAA TTT CTG GTT CAA TGA AGA
Val Asn Leu Gln Lys Thr Lys Pro Gln Val Lys Asp Gln Val Thr Ser>

1590        1600        1610        1620        1630
TTG AAT GAA GAT TTG ACT ACT ATG TCT ATA GAT TCC AGT AGT CCT GTA
AAC TTA CTT CTA AAC TGA TGA TAC AGA TAT CTA AGG TCA TCA GGA CAT
Leu Asn Glu Asp Leu Thr Thr Met Ser Ile Asp Ser Ser Ser Pro Val>

1640        1650        1660        1670        1680
TTT TTA GAG GTT ATT GAT CCA ATT ACA AAT TTA GGA ACT CTT CAA CTT
AAA AAT CTC CAA TAA CTA GGT TAA TGT TTA AAT CCT TGA GAA GTT GAA
Phe Leu Glu Val Ile Asp Pro Ile Thr Asn Leu Gly Thr Leu Gln Leu>
```

FIGURE 16 (4 of 5)

```
          1690        1700        1710        1720
           *           *           *           *
ATT GAT TTA AAT ACT GGT GTT AGG CTT AAA GAA AGC ACT CAG CAA GGC
TAA CTA AAT TTA TGA CCA CAA TCC GAA TTT CTT TCG TGA GTC GTT CCG
Ile Asp Leu Asn Thr Gly Val Arg Leu Lys Glu Ser Thr Gln Gln Gly>

1730        1740        1750        1760        1770
 *           *           *           *           *
ATT CAG CGG TAT GGA ATT TAT GAA CGT GAA AAA GAT TTG GTT GTT ATT
TAA GTC GCC ATA CCT TAA ATA CTT GCA CTT TTT CTA AAC CAA CAA TAA
Ile Gln Arg Tyr Gly Ile Tyr Glu Arg Glu Lys Asp Leu Val Val Ile>

 1780        1790        1800        1810        1820
  *           *           *           *           *
AAA ATG GAT TCA GGA AAA GCT AAG CTT CAG ATA CTT GAT AAA CTT GAA
TTT TAC CTA AGT CCT TTT CGA TTC GAA GTC TAT GAA CTA TTT GAA CTT
Lys Met Asp Ser Gly Lys Ala Lys Leu Gln Ile Leu Asp Lys Leu Glu>

   1830        1840        1850        1860        1870
    *           *           *           *           *
AAT TTA AAA GTG GTA TCA GAG TCT AAT TTT GAG ATT AAT AAA AAT TCA
TTA AAT TTT CAC CAT AGT CTC AGA TTA AAA CTC TAA TTA TTT TTA AGT
Asn Leu Lys Val Val Ser Glu Ser Asn Phe Glu Ile Asn Lys Asn Ser>

      1880        1890        1900        1910        1920
       *           *           *           *           *
TCT CTT TAT GTT GAT TCT AAA ATG ATT TTA GTA GCT GTT AGG GAT AAA
AGA GAA ATA CAA CTA AGA TTT TAC TAA AAT CAT CGA CAA TCC CTA TTT
Ser Leu Tyr Val Asp Ser Lys Met Ile Leu Val Ala Val Arg Asp Lys>

          1930        1940        1950        1960
           *           *           *           *
GAT AGT AGT AAT GAT TGG AGA TTG GCC AAA TTT TCT CCT AAA AAT TTA
CTA TCA TCA TTA CTA ACC TCT AAC CGG TTT AAA AGA GGA TTT TTA AAT
Asp Ser Ser Asn Asp Trp Arg Leu Ala Lys Phe Ser Pro Lys Asn Leu>

1970        1980        1990        2000        2010
 *           *           *           *           *
GAT GAG TTT ATT CTT TCA GAG AAT AAA ATT ATG CCT TTT ACT AGC TTT
CTA CTC AAA TAA GAA AGT CTC TTA TTT TAA TAC GGA AAA TGA TCG AAA
Asp Glu Phe Ile Leu Ser Glu Asn Lys Ile Met Pro Phe Thr Ser Phe>

 2020        2030        2040        2050        2060
  *           *           *           *           *
TCT GTG AGA AAA AAT TTT ATT TAT TTG CAA GAT GAG TTT AAA AGT CTA
AGA CAC TCT TTT TTA AAA TAA ATA AAC GTT CTA CTC AAA TTT TCA GAT
Ser Val Arg Lys Asn Phe Ile Tyr Leu Gln Asp Glu Phe Lys Ser Leu>

   2070        2080        2090        2100
    *           *           *           *
GTT ATT TTA GAT GTA AAT ACT TTA AAA AAA GTT AAG TA
CAA TAA AAT CTA CAT TTA TGA AAT TTT TTT CAA TTC AT
Val Ile Leu Asp Val Asn Thr Leu Lys Lys Val Lys Xxx>
```

FIGURE 16 (5 of 5)

p93 - K48

```
   1 ATGAAAAAAT TGTTACTAAT CTTTAGTTTT TTTCTTATTT CTTTGAATGG ATTTCCTCTT
  61 AATTCAAGGG AAGTTGATAA GGAAAAATTA AAGGATTTTG TTAATATGGA TCTTGAGTTT
 121 GTAAACTATA AAGGTCCTTA TGATTCTACA AATACATATG AACAAATAGT AGGTATTGGT
 181 GAGTTTTTAG CAAGACCATT GATTAATTCC AATAGCAACT CAATTTATTA TGGTAAATAT
 241 TTTATTAATA GATTTATTGA TGATCAAGAT AAAAAAGCAA GCGTTGATGT TTTTTCTATT
 301 GGTAGTAGGT CACAGCTTGA CAGTATATTG AATCTAAGAA GAATTCTTAC AGGGTATTTG
 361 ATAAAGTCTT TTGATTATGA AAGATCTAGT GCTGAATTAA TTGCTAAGGT TATTACAATA
 421 CATAATGCTG TTTATAGAGG GGATTTAAAT TATTATAAAG AGGTTTATAT TGAGGCTGCT
 481 TTAAAGTCTT TAACTAAAGA AAATGCAGGT CTTTCTAGAG TGTACAGTCA ATGGGCTGGA
 541 AAGACACAAA TATTTATTCC TCTTAAAAAG AATATTTTAT CTGGAAAAGT TGAGTCTGAC
 601 ATTGATATTG ACAGTTTGGT TACAGATAAG GTTGTGGCAG CTCTTTTAAG CGAGAATGAA
 661 GCAGGTGTTA ACTTTGCAAG AGATATTACA GATATTCAAG GCGAAACTCA TAAAGCAGAT
 721 CAAGATAAAA TTGATATTGA ATTAGATAAT GTTCATAAAA GTGATTCCAA TATAACAGAG
 781 ACTATTGAGA ATTTAAGAGA TCAGCTTGAA AAGGCTACAG ATGAAGAGCA TAGAAAAGAG
 841 ATTGAAAGTC AGGTTGATGC TAAAAAGAAA CAAAAAGAAG AACTAGATAA AAAGGCAATC
 901 GATCTTGATA AAGCCCAACA AAAATTAGAT TCTTCTGAAG ATAATTTAGA TATTCAAAGG
 961 GATACTGTTA GAGAGAAGAT TCAAGAGGAT ATTGACGAGA TTAATAAAGA AAAGAATTTG
1021 CCAAAACCTG GTGATGTAAG TTCTCCTAAA GTTGATAAGC AGCTACAAAT AAAAGAGAGT
1081 CTAGAAGACT TGCAGGAACA GCTTAAAGAA ACTAGCGATG AAAATCAAAA AAGAGAAATT
1141 GAAAAGCAAA TTGAAATCAA AAAAAGTGAT GAAGAACTTT TAAAAAGTAA AGATCCTAAA
1201 GCATTAGATC TTAATGGAGA TTTAAATTCT AAAGTTTCTA GTAAAGAAAA AATTAAAGGC
1261 AAAGAAGGAG AAATAGTCAA AGAGGAATCA AAGGCAAGTT TAGCTGATTT GAATAATGAC
1321 GAAAATCTTA TGAGGCCGGA AGATCAAAAA TTATCTGAGG ATAAAAAATT AGATAGTAAA
1381 AAAAATTTAA AACCTGTTTC TGAGATTGAG AGAGTAAATG AAATTTCGAA GTCTAACAAC
1441 AATGAGATTA GTGAATCATC ACCATTATAT AAGCCTTCTT ATAGCGATAT GGATTCAAAA
1501 GAGGGTATAG ATAATAAAGA TGTTAACTTG CAAGAAACCA AGTCTCAAAC TAAAAGTCAA
1561 CCTACTTCTT TAAATCAAGA TTTGACTACT ATGTCTATAG ATTCTAGTAA TCCTGTATTT
1621 TTAGAGGTTA TTGATCCTAT TACAAATTTA GGAACGCTTC AACTTATTGA TTTGAATACC
1681 GGTGTTAGAC TTAAAGAAAG TACTCAGCAA GGCATTCAGC GGTATGGAAT TTATGAACGT
1741 GAAAAAGATT TAGTTGTTAT TAAAATGGAT TCAGGAAAAG CCAAGCTTCA AATACTTAAT
1801 AAACTTGAGA ATTTAAAAGT GATATCGGAG TCTAATTTTG AGATTAATAA AAATTCATCT
1861 CTTTATGTTG ACTCTAAAAT GATTTTAGTA GTTGTGAGAG ATAGTGGTAA TGTTTGGAGA
1921 TTGGCTAAAT TTTCTCCTAA AAATTTAAAT GAGTTTATTC TTTCAGAGAA TAAAATTTTG
1981 CCTTTTACTA GCTTTTCTGT GAGAAAGAAT TTTATTTATT TGCAGGATGA GTTTAAAAGT
2041 CTTATTACTT TAGATGTAAA TACTTTAAAA AAAGTTAAGT A
```

FIGURE 17

p93 - BO

```
   1 ATGAAAAAAA TGTTACTAAT CTTTAGTTTT TTTCTTGTTT TTTTAAATGG ATTTCCTCTT
  61 AATGCAAGGG AAGTTGATAA GGAAAAATTA AAGGACTTTG TTAATATGGA TCTTGAATTT
 121 GTTAATTACA AGGGTCCTTA TGATTCTACA GATACATATG AACAAATAGT AGGTATTGGG
 181 GAGTTTTTAG CAAGGCCGTT GAACAATTCC AATAGTAATT CAAGTTATTA TGGTAAATAT
 241 TTTGTTAATA GATTTATTGA CGATCAAGAT AAAAAAGCAA GTGTTGATAT TTTTTCTATT
 301 GGTAGTAAGT CAGAGCTTGA TAGTATATTA AATCTAAGAA GAATTCTTAC AGGGTATTTA
 361 ATGAAGTCTT TTGATTATGA GAGGTCTAGT GCGGAATTAA TTGCTAAAGC TATTACAATA
 421 TATAATGCTG TTTATAGAGG AGATTTAGAT TATTACAAAG AGTTTTATAT TGAGGCTTCT
 481 TTGAAGTCTT TGACTAAAGA AAATGCAGGT CTTTCTAGGG TGTACAGTCA ATGGGCTGGG
 541 AAGACACAAA TATTTATTCC TCTTAAAAAG AATATTTTAT CTGGAAATGT TGAGTCTGAC
 601 ATTGATATTG ATAGTTTGGT TACAGATAAG GTGGTGGCAG CTCTTTTAAG TGAGAATGAA
 661 TCAGGTGTTA ACTTTGCAAG AGATATTACA GACATTCAAG GCGAAACTCA TAAAGCAGAT
 721 CAAGATAAAA TTGATATTGA ATTAGATAAT TTTCATGAAA GTGATTCCAA TATAACAGAA
 781 ACTATTGAGA ATTTAAGGGA TCAGCTTGAA AAAGCTACAG ATGAAGAGCA TAAAAAAGAG
 841 ATTGAAAGTC AGGTTGATGC TAAAAAGAAA CAAAAGGAAG AATTAGATAA AAAGGCAATT
 901 GATCTTGATA AAGCTCAACA AAAATTAGAT TTTGCTGAAG ATAATCTAGA TATTCAAAGG
 961 GATACTGTTA GAGAGAAGCT TCAAGAAAAT ATTAACGAGA CTAATAAGGA AAAGAATTTA
1021 CCAAAGCCTG GTGATGTAAG TTCTCCTAAG GTTGATAAGC AGTTGCAGAT AAAAGAGAGT
1081 CTAGAAGATT TGCAAGAGCA GCTTAAAGAA GCTAGTGATG AAAATCAAAA AAGAGAAATA
1141 GAAAAGCAAA TTGAAATCAA AAAAAATGAT GAAGAACTTT TTAAAAATAA AGATCATAAA
1201 GCATTAGATC TTAAGCAAGA ATTAAATTCT AAAGCTTCTA GTAAAGAAAA AATTGAAGGC
1261 GAAGAAGAGG ATAAAGAATT AGATAGTAAA AAAAATTTAG AGCCTGTTTC TGAGGCTGAT
1321 AAAGTAGATA AAATTTCCAA GTCTAACAAC AATGAGGTTA GTAAATTATC CCCGTTAGAT
1381 GAGCCTTCTT ATAGCGACAT TGATTCGAAA GAGGGTGTAG ATAACAAAGA TGTTGATTTG
1441 CAAAAAACTA AACCCCAAGT TGAAAGTCAA CCTACTTCGT TAAATGAAGA TTTGATTGAT
1501 GTGTCTATAG ATTCCAGTAA TCCTGTCTTT TTAGAGGTTA TCGATCCGAT TACAAATTTA
1561 GGAACGCTTC AACTTATTGA TTTGAATACC GGTGTTAGAC TTAAAGAAAG TGCTCAACAA
1621 GGTATTCAGC GATATGGAAT TTATGAACGT GAAAAAGATT TGGTTGTTAT TAAAATAGAT
1681 TCAGGAAAAG CTAAGCTTCA GATACTTGAT AAACTCGAGA ATTTAAAAGT GATATCAGAG
1741 TCTAATTTTG AGATTAATAA AAATTCATCT CTTTATGTTG ACTCTAGAAT GATTTTAGTA
1801 GTTGTTAAGG ACGATAGTAA TGCTTGGAGA TTGGCTAAAT TTTCTCCTAA AAATTTAGAT
1861 GAATTTATTC TGTCAGAAAA TAAAATTTTG CCTTTTACTA GCTTTGCTGT GAGAAAGAAT
1921 TTTATTTATT TGCAAGATGA ACTTAAAAGC TTAGTTACTT TAGATGTAAA TACTTTAAAA
1981 AAAGTTAAGT A
```

FIGURE 18

p93 - pIRO

```
   1 ATGAAAAAAA TGTTACTAAT CTTTAGTTTT TTTCTTATTT CTTTGAATGG ATTTCCCCTT
  61 AATGCAAGGG AAGTTGATAA GGAAAAATTA AAGGACTTTG TTAATATGGA TCTTGAGTTT
 121 GTAAACTATA AAGGTCCTTA TGATTCTACA AATACATATG AACAAATAGT AGGTATTGGT
 181 GAGTTTTTAG CAAGACCATT GATTAATTTC AATAGCAACT CAAGTTATTA TGGTAAATAT
 241 TTTATTAATA GATTTATTGA CGATCAAGAT AAAAAAGCAA GCGTTGATGT TTTTTCTATT
 301 AGTAGTAAGT CACAGCTTGA CAGTATATTG AATTTAAGAA GAATTCTTAC AGGGTATTTG
 361 ATAAAGTCTT TTGATTATGA AAGATCTAGT GCTGAATTAA TTGCCAAGGT TATTACAATA
 421 CATAATGCTG TTTATAGAGG TGATTTAAAT TATTATAAAG AGTTTTATAT TGAGTCTGCT
 481 TTAAAGTCTT TAACTAAAGA AAATGCAGGT CTTTCTAGAG TGTACAGTCA ATGGGCTGGA
 541 AAGACACAAA TATTTATTCC TCTTAAAAAG AATATTTTAT CTGGAAAAAT TGAGTCTGAC
 601 ATTGATATTG ATAGTTTGGT TACAGATAAG GTTGTGGCAG GTGTTTTAAG CGAAAATGAA
 661 GCAGGTGTTA ACTTTGCAAG GGATATTACA GATATTCAAG GAGAAACTCA TAAAGCAGAT
 721 CAAGATAAAA TTGATATTGA ATTAGATAAT GTTCATGAAA GTGATTCCAA TATAACAGAA
 781 ACTATTGAGA ATTTAAGAGA TCAGCTTGAA AAGGCTACAG ATGAAGAGCA TAGAAAAGAG
 841 ATTGAAAGTC AAGTTGATGC TAAAAAGAAA CAAAAAGAAG AACTAGATAA AAAGGCAATC
 901 GATCTTGATA AAGCCCAACA AAAATTAGAT TTTTCTGAAG ATAATTTAGA TATTCAAAGG
 961 GATACTGTTA GAGAGAAGAT TCAAGAGGAT ATTAACGAGA TTAATAAGGA AAAGAATTTA
1021 CCAAAACCTG GTGATGTAAG TTCTCCTAAA GTTGATAAGC AGCTACAAAT AAAAGAGAGT
1081 CTAGAAGACT TGCAGGAGCA GCTTAAAGAA ACTAGCGATG AAAATCAAAA AAGAGAAATT
1141 GAAAAGCAAA TTGAAATCAA AAAAAGTGAT GAAGAACTTT TAAAAAGCAA AGATCCTAAA
1201 GCATTAGATC TTAATCGAGA TTTAAATTCT AAAGCTTCTA GTAAAGAAAA AATTAAAGGC
1261 AAAGAAAAAG AAATAGTCAA AGAGAAATCA AAGGTAAGTT TAGGTGATTT GGATAATGAC
1321 GAAACCCTTA TGACGCCGGA AGATCAAAAA TTATCTGAGG ATAAAAAATT AGATAGTAAA
1381 AAAAATTTAA AACCTGTTTC TGAGATTGAG AGAGTAAATG AAATTTCAAA GTCTAACAAC
1441 AATGAGGTTA GCAAATCATC ACCATTAGAT AAGCCTTCTT ATAGTGATAT CGATTCAAAA
1501 GAGGTTGTAG ATAATAAAGA TGTTAATTTG CAAGAAACCA AGCCTCAAGC TAAAAGTCAA
1561 TCTACTTCTT TAAATCAAGA TTTGATTACT ATGTCTATAG ATTCTAGTAA TCCTGTATTT
1621 TTAGAGGTTA TTGATCCTAT TACAAATTTA GGAATGCTTC AACTTATTGA TTTAAATACT
1681 GGTGTTAGAC TTAAAGAAAG CACTCAGCAA GGCATTCAGC GTTATGGAAT TTATGAACGT
1741 GAAAAAGATT TAGTTGTTAT TAAAATGGAT TCAGGAAAAG CTAAGCTTCA AATACTTAAT
1801 AAACTTGAGA ATTTAAAAGT GATATCAGAG TCTAATTTTG AGATTAATAA AAATTCATCT
1861 CTTTATGTTG ACTCTAAAAT GATTTTAGTA GCTGTGAAAG ATAGTGGTAA TGTTTGGAGA
1921 TTGGCTAAAT TTTCTCCTAA AAATTTAGAT GAGTTTATTC TTTCAGAGAA TAAAATTTTG
1981 CCTTTTACTA GCTTTTCTGT GAGAAAGAAT TTTATTTATT TGCAAGATGA GTTTAAAAGT
2041 CTTATTACTT TAGATGTAAA TACTTTAAAA AAAGTTAAGT A
```

FIGURE 19

p93 - pGau

```
   1 ATGAAAAAAA TGTTACTAAT CTTTAGTTTT TTTCTTGTTT TTTTAAATGG ATTTCCTCTT
  61 AATGCAAGGG AAGTTGATAA GGAAAAATTA AAGGACTTTG TTAATATGGA TCTTGAATTT
 121 GTTAATTACA AGGGTCCTTA TGATTCTACA AATACATATG AACAAATAGT AGGTATTGGG
 181 GAGTTTTTAG CAAGGCCGTT GATCAATTCC AATAGTAATT CAAGTTATTA TGGTAAATAT
 241 TTTGTTAATA GATTTATTGA CGATCAAGAT AAAAAAGCAA GTGTTGATAT TTTTTCTATT
 301 GGTAGTAAGT CAGAGCTTGA TAGTATATTA AATCTAAGAA GAATTCTTAC AGGGTATTTA
 361 ATGAAGTCTT TTGATTATGA GAGGTCTAGT GCGGAATTAA TTGCTAAAGC TATTACAATA
 421 TATAATGCTG TTTATAGAGG AGATTTAGAT TATTACAAAG AGTTTTATAT TGAGGCTTCT
 481 TTGAAGTCTT TGACTAAAGA AAATGCAGGT CTTTCTAGGG TGTACAGTCA ATGGGCTGGG
 541 AAGACACAAA TATTTATTCC TCTTAAAAAG AATATTTTAT CTGGAAATGT TGAGTCTGAC
 601 ATTGATATTG ATAGTTTGGT TACAGATAAG GTGGTGGCAG CTCTTTTAAG TGAGAATGAA
 661 TCAGGTGTTA ACTTTGCAAG AGATATTACA GACATTCAAG GCGAAACTCA TAAAGCAGAT
 721 CAAGATAAAA TTGATATTGA ATTAGATAAT ATTCATGAAA GTGATTCCAA TATAACAGAA
 781 ACTATTGAGA ATTTAAGGGA TCAGCTTGAA AAAGCTACAG ATGAAGAGCA TAAAAAAGAG
 841 ATTGAAAGTC AGGTTGATGC TAAAAAGAAA CAAAAGGAAG AATTAGATAA AAAGGCAATT
 901 GATCTTGATA AAGCTCAACA AAAATTAGAT TTTGCTGAAG ATAATCTAGA TATTCAAAGG
 961 GATACTGTTA GAGAGAAGCT TCAAGAGAAT ATTAACGAGA CTAATAAGGA AAAGAATTTA
1021 CCAAAGCCTG GTGATGTAAG TTCTCCTAAA GTTGATAAGC AACTACAAAT AAAAGAGAGC
1081 CTGGAAGATT TGCAGGAGCA GCTTAAAGAA ACTGGTGATG AAAATCAGAA AAGAGAAATT
1141 GAAAAGCAAA TTGAAATCAA AAAAAGTGAT GAAAAGCTTT TAAAAAGTAA AGATGATAAA
1201 GCAAGTAAAG ATGGTAAAGC CTTGGATCTT GATCGAGAAT TAAATTCTAA AGCTTCTAGC
1261 AAAGAAAAAA GTAAAGCCAA GGAAGAAGAA ATAACCAAGG GTAAGTCACA GAAAAGCTTA
1321 GGCGATTTGA ATAATGATGA AAATCTTATG ATGCCAGAAG ATCAAAAATT ACCTGAGGTT
1381 AAAAAATTAG ATAGCAAAAA AGAATTTAAA CCTGTTTCTG AGGTTGAGAA ATTAGATAAG
1441 ATTTTCAAGT CTAATAACAA TGTTGGAGAA TTATCACCGT TAGATAAATC TTCTTATAAA
1501 GACATTGATT CAAAAGAGGA GACAGTTAAT AAAGATGTTA ATTTGCAAAA GACTAAGCCT
1561 CAGGTTAAAG ACCAAGTTAC TTCTTTGAAT GAAGATTTGA CTACTATGTC TATAGATTCC
1621 AGTAGTCCTG TATTTTTAGA GGTTATTGAT CCAATTACAA ATTTAGGAAC TCTTCAACTT
1681 ATTGATTTAA ATACTGGTGT TAGGCTTAAA GAAAGCACTC AGCAAGGCAT TCAGCGGTAT
1741 GGAATTTATG AACGTGAAAA AGATTTGGTT GTTATTAAAA TGGATTCAGG AAAAGCTAAG
1801 CTTCAGATAC TTGATAAACT TGAAAATTTA AAAGTGGTAT CAGAGTCTAA TTTTGAGATT
1861 AATAAAAATT CATCTCTTTA TGTTGATTCT AAAATGATTT TAGTAGCTGT TAGGGATAAA
1921 GATAGTAGTA ATGATTGGAG ATTGGCCAAA TTTTCTCCTA AAAATTTAGA TGAGTTTATT
1981 CTTTCAGAGA ATAAAATTAT GCCTTTTACT AGCTTTTCTG TGAGAAAAAA TTTTATTTAT
2041 TTGCAAGATG AGTTTAAAAG TCTAGTTATT TTAGATGTAA ATACTTTAAA AAAAGTTAAG
2101 TAAAGCC
```

FIGURE 20

p93 - pKO

```
   1 ATGAAAAAAA TGTTACTAAT CTTTAGTTTT TTTCTTGTTT TTTTAAATGG ATTTCCTCTT
  61 AATGCAAGGG AAGTTGATAA GGAAAAATTA AAGGACTTTG TTAATATGGA TCTTGAATTT
 121 GTTAATTACA AGGGTCCTTA TGATTCTACA AATACATATG AACAAATAGT AGGTATTGGG
 181 GAGTTTTTAG CAAGGCCGTT GATCAATTCC AATAGTAATT CAAGTTATTA TGGTAAATAT
 241 TTTGTTAATA GATTTATTGA CGATCAAGAT AAAAAAGCAA GTGTTGATAT TTTTTCTATT
 301 GGTAGTAAGT CAGAGCTTGA TAGTATATTA AATCTAAGAA GAATTCTTAC AGGGTATTTA
 361 ATGAAGTCTT TTGATTATGA GAGGTCTAGT GCGGAATTAA TTGCTAAAGC TATTACAATA
 421 TATAATGCTG TTTATAGAGG AGATTTAGAT TATTACAAAG AGTTTTATAT TGAGGCTTCT
 481 TTGAAGTCTT TGACTAAAGA AAATGCAGGT CTTTCTAGGG TGTACAGTCA ATGGGCTGGG
 541 AAGACACAAA TATTTATTCC TCTTAAAAAG AATATTTTAT CTGGAAATGT TGAGTCTGAC
 601 ATTGATATTG ATAGTTTGGT TACAGATAAG GTGGTGGCAG CTCTTTTAAG TGAGAATGAA
 661 TCAGGTGTTA ACTTTGCAAG AGATATTACA GACATTCAAG GCGAAACTCA TAAAGCAGAT
 721 CAAGATAAAA TTGATATTGA ATTAGATAAT TTTCATGAAA GTGATTCCAA TATAACAGAA
 781 ACTATTGAGA ATTTAAGGGA TCAGCTTGAA AAAGCTACAG ATGAAGAGCA TAAAAAAGAG
 841 ATTGAAAGTC AGGTTGATGC TAAAAAGAAA CAAAAGGAAG AATTAGATAA AAAGGCAATT
 901 GATCTTGATA AAGCTCAACA AAAATTAGAT TTTGCTGAAG ATAATCTAGA TATTCAAAGG
 961 GATACTGTTA GAGAGAAGCT TCAAGAAAAT ATTAACGAGA CTAATAAGGA AAAGAATTTA
1021 CCAAAGCCTG GTGATGTAAG TTCTCCTAAG GTTGATAAGC AGTTGCAGAT AAAAGAGAGT
1081 CTAGAAGATT TGCAAGAGCA GCTTAAAGAA GCTAGTGATG AAAATCAAAA AAGAGAAATA
1141 GAAAAGCAAA TTGAAATCAA AAAAAATGAT GAAGAACTTT TTAAAAATAA AGATCATAAA
1201 GCATTAGATC TTAAGCAAGA ATTAAATTCT AAAGCTTCTA GTAAAGAAAA AATTGAAGGC
1261 GAAGAAGAGG ATAAAGAATT AGATAGTAAA AAAAATTTAG AGCCTGTTTC TGAGGCTGAT
1321 AAAGTAGATA AAATTTCCAA GTCTAACAAC AATGAGGTTA GTAAATTATC CCCGTTAGAT
1381 GAGCCTTCTT ATAGCGACAT TGATTCGAAA GAGGGTGTAG ATAACAAAGA TGTTGATTTG
1441 CAAAAAACTA AACCCCAAGT TGAAAGTCAA CCTACTTCGT TAAATGAAGA CTTGATTGAT
1501 GTGTCTATAG ATTCCAGTAA TCCTGTCTTT TTAGAGGTTA TCGATCCGAT TACAAATTTA
1561 GGAACGCTTC AACTTATTGA TTTGAATACC GGTGTTAGAC TTAAAGAAAG TGCTCAACAA
1621 GGTATTCAGC GATATGGAAT TTATGAACGT GAAAAAGATT TGGTTGTTAT TAAAATAGAT
1681 TCAGGAAAAG CTAAGCTTCA GATACTTGAT AAACTCGAGA ATTTAAAAGT GATATCAGAG
1741 TCTAATTTTG AGATTAATAA AAATTCATCT CTTTATGTTG ACTCTAGAAT GATTTTAGTA
1801 GTTGTTAAGG ACGATAGTAA TGCTTGGAGA TTGGCTAAAT TTTCTCCTAA AAATTTAGAT
1861 GAATTTATTC TGTCAGAAAA TAAAATTTTG CCTTTTACTA GCTTTGCTGT GAGAAAGAAT
1921 TTTATTTATT TGCAAGATGA ACTTAAAAGC TTAGTTACTT TAGATGTAAA TACTTTAAAA
1981 AAAGTTAAGT A
```

FIGURE 21

p93 - 25015

```
   1 ATGAAAAAAA TGTTACTAAT CTTTAGTTTT TTTCTTATTT TTTTGAATGG ATTTCCTCTT
  61 AATGCAAGGA AAGTTGATAA GGAAAAATTA AAGGATTTTG TTAATATGGA TCTTGAGTTT
 121 GTAAATTATA AAGGTCCTTA TGATTCTACA AATACGTATG AACAAATAGT GGGTATTGGG
 181 GAGTTTTTAG CAAGACCGCT GACCAATTCC AATAGCAACT CAAGTTATTA TGGCAAATAT
 241 TTTATTAATA GATTTATTGA TGATCAAGAT AAAAAAGCAA GTGTTGATGT TTTTTCTATA
 301 AGCAGCAAAT CAGAGCTTGA CAGTATATTG AATTTAAGAA GAATTCTTAC AGGGTATATA
 361 ATAAAGTCTT TCGATTATGA CAGGTCTAGT GCAGAATTAA TTGCTAAGGT TATTACAATA
 421 TATAATGCTG TTTATAGAGG AGATTTGGAT TATTATAAAG GGTTTTATAT TGAGCCTGCT
 481 TTGAAGTCTT TAACTAAAGA AAACGCAGGT CTTTCTAGGG TTTACAGTCA GTGGGCTGGA
 541 AAGACTCAAA TATTTATTCC TCTTAAAAAG GATATTTTGT CTGGAAATAT TGAATCTGAC
 601 ATTGATATTG ACAGTTTGGT TACAGATAAG GTGATAGCAG CTCTTTTAAG CGAAAATGAA
 661 GCAGGCGTTA ACTTTGCAAG AGATATTACA GATATTCAAG GCGAAACTCA TAAGGCAGAT
 721 CAAGATAAGA TTGATACTGA ATTAGACAAT ATCCATGAAA GCGATTCTAA TATAACAGAA
 781 ACTATTGAAA ATTTAAGGGA TCAGCTTGAA AAAGCTACAG ATGAAGAGCA TAAAAAAGAG
 841 ATTGAAAGTC AGGTTGATGC TAAAAAGAAA GAAAAGGAAG AGCTAGATAA AAAGGCAATC
 901 AATCTTGATA AAGCTCAGCA AAAATTAGAC TCTGCTGAAG ATAATTTAGA TGTTCAAAGA
 961 GATACTGTTA GAGAGAAAAT TCAAGAGGAT ATTAATGAGA TTAATAAGGA AAAGAATTTG
1021 CCAAAACCTG GTGATGTAAG TTCTCCTAAA GTTGATAAGC AACTGCAAAT AAAAGAGAGT
1081 CTAGAAGATT TGCAGGAGCA GCTTAAAGAA GCTGGTGATG AAAATCAGAA AAGAGAAATT
1141 GAGAAGCAAA TTGAAATCAA AAAAAGGGAC GAAGAACTTT TAAAAAGTAA AGATGGCAAA
1201 GTAAGTAAAG ATTATGAAGC ATTAGATCTT GATCGAGAAT TATCCAAAGC TTCTAGTAAA
1261 GAAAAAAGTA AGGTCAAGGA AGAAGAAATA ACTAAAGGTA AATCACGGGC AAGCTTAGGC
1321 GATTTGAATA ATGATAAAAA CCTTATGTTG CCAGAAGATC AAAAATTACC TGAAGATAAA
1381 AAATTGGATA GTAAATTAGA TGGTAAAAAA GAATTTAAAC CAGTTTCTGA GGTTGAAAAA
1441 TTAGATAAGA TTTCCAAGTC TAATAACAAT GAGGTTGGCA AGTTATCACC ATTAGATAAG
1501 CCTTCTTATG ATGATATTGA TTCAAAAGAG GAGGTAGATA ATAAAGCTAT TAATTTGCAA
1561 AAGATCGACC CTAAAGTTAA AGACCAAACT ACTTCTTTGA ATGAAGATTT GGATAAAGAT
1621 TTGACTACTA TGTCTATAGA TTCCAGCAGT CCTGTATTTC TAGAGGTTAT TGATCCTATT
1681 ACAAATTTAG GAACCCTGCA GCTTATTGAT TTAAATACTG GGGTTAGGCT TAAGGAAAGC
1741 ACTCAGCAAG GCATTCAGCG GTATGGAATT TATGAACGTG AAAAAGATTT GGTTGTTATT
1801 AAAATGGATT CAGGAAAGGC TAAGCTTCAA ATACTTAATA AGCTTGAAAA TTTGAAAGTG
1861 GTATCAGAGT CTAATTTTGA GATCAATAAA AATTCATCTC TTTATGTTGA CTCTAAAATG
1921 ATTTTGGCAG CTGTTAGAGA TAAGGATGAT AGCAATGCTT GGAGATTGGC TAAATTTTCT
1981 CCTAAAAATT TGGATGAGTT TATTCTTTCA GAGAATAAAA TTTTGCCTTT TACTAGCTTT
2041 TCTGTGAGAA AAAATTTTAT TTATTTGCAA GATGAGCTTA AAAATCTAGT TATTTTAGAT
2101 GTAAATACTT TAAAAAAAGT TAAGTA
```

FIGURE 22

K48 OSP A/ PGAU OSP A FUSION


```
          10          20          30          40
     *    *    *    *    *    *    *    *    *
ATG AAA AAA TAT TTA TTG GGA ATA GGT CTA ATA TTA GCC TTA ATA GCA
TAC TTT TTT ATA AAT AAC CCT TAT CCA GAT TAT AAT CGG AAT TAT CGT
Met Lys Lys Tyr Leu Leu Gly Ile Gly Leu Ile Leu Ala Leu Ile Ala>

50          60          70          80          90
 *    *    *    *    *    *    *    *    *    *
TGT AAG CAA AAT GTT AGC AGC CTT GAT GAA AAA AAT AGC GTT TCA GTA
ACA TTC GTT TTA CAA TCG TCG GAA CTA CTT TTT TTA TCG CAA AGT CAT
Cys Lys Gln Asn Val Ser Ser Leu Asp Glu Lys Asn Ser Val Ser Val>

  100         110         120         130         140
   *    *    *    *    *    *    *    *    *    *
GAT TTA CCT GGT GGA ATG ACA GTT CTT GTA AGT AAA GAA AAA GAC AAA
CTA AAT GGA CCA CCT TAC TGT CAA GAA CAT TCA TTT CTT TTT CTG TTT
Asp Leu Pro Gly Gly Met Thr Val Leu Val Ser Lys Glu Lys Asp Lys>

    150         160         170         180         190
 *    *    *    *    *    *    *    *    *    *    *
GAC GGT AAA TAC AGT CTA GAG GCA ACA GTA GAC AAG CTT GAG CTT AAA
CTG CCA TTT ATG TCA GAT CTC CGT TGT CAT CTG TTC GAA CTC GAA TTT
Asp Gly Lys Tyr Ser Leu Glu Ala Thr Val Asp Lys Leu Glu Leu Lys>

       200         210         220         230         240
   *    *    *    *    *    *    *    *    *    *    *
GGA ACT TCT GAT AAA AAC AAC GGT TCT GGA ACA CTT GAA GGT GAA AAA
CCT TGA AGA CTA TTT TTG TTG CCA AGA CCT TGT GAA CTT CCA CTT TTT
Gly Thr Ser Asp Lys Asn Asn Gly Ser Gly Thr Leu Glu Gly Glu Lys>

          250         260         270         280
     *    *    *    *    *    *    *    *    *
ACT GAC AAA AGT AAA GTA AAA TTA ACA ATT GCT GAT GAC CTA AGT CAA
TGA CTG TTT TCA TTT CAT TTT AAT TGT TAA CGA CTA CTG GAT TCA GTT
Thr Asp Lys Ser Lys Val Lys Leu Thr Ile Ala Asp Asp Leu Ser Gln>

290         300         310         320         330
 *    *    *    *    *    *    *    *    *    *
ACT AAA TTT GAA ATT TTC AAA GAA GAT GCC AAA ACA TTA GTA TCA AAA
TGA TTT AAA CTT TAA AAG TTT CTT CTA CGG TTT TGT AAT CAT AGT TTT
Thr Lys Phe Glu Ile Phe Lys Glu Asp Ala Lys Thr Leu Val Ser Lys>

  340         350         360         370         380
   *    *    *    *    *    *    *    *    *    *
AAA GTA ACC CTT AAA GAC AAG TCA TCA ACA GAA GAA AAA TTC AAC GAA
TTT CAT TGG GAA TTT CTG TTC AGT AGT TGT CTT CTT TTT AAG TTG CTT
Lys Val Thr Leu Lys Asp Lys Ser Ser Thr Glu Glu Lys Phe Asn Glu>
```

FIGURE 23 (1 of 3)

K48 OSP A/ PGAU OSPA FUSION

```
    390          400          410          420          430
    .    .       .    .       .    .       .    .       .    .
AAG GGT GAA ACA TCT GAA AAA ACA ATA GTA AGA GCA AAT GGA ACC AGA
TTC CCA CTT TGT AGA CTT TTT TGT TAT CAT TCT CGT TTA CCT TGG TCT
Lys Gly Glu Thr Ser Glu Lys Thr Ile Val Arg Ala Asn Gly Thr Arg>

       440          450          460          470          480
  .    .       .    .       .    .       .    .       .    .
CTT GAA TAC ACA GAC ATA AAA AGC GAT GGA TCC GGA AAA GCT AAA GAA
GAA CTT ATG TGT CTG TAT TTT TCG CTA CCT AGG CCT TTT CGA TTT CTT
Leu Glu Tyr Thr Asp Ile Lys Ser Asp Gly Ser Gly Lys Ala Lys Glu>

          490          500          510          520
     .    .       .    .       .    .       .    .       .
GTT TTA AAA GAC TTT ACT CTT GAA GGA ACT CTA GCT GCT GAC GGC AAA
CAA AAT TTT CTG AAA TGA GAA CTT CCT TGA GAT CGA CGA CTG CCG TTT
Val Leu Lys Asp Phe Thr Leu Glu Gly Thr Leu Ala Ala Asp Gly Lys>

530          540          550          560          570
  .    .       .    .       .    .       .    .       .    .
ACA ACA TTG AAA GTT ACA GAA GGC ACT GTT GTT TTA AGC AAG AAC ATT
TGT TGT AAC TTT CAA TGT CTT CCG TGA CAA CAA AAT TCG TTC TTG TAA
Thr Thr Leu Lys Val Thr Glu Gly Thr Val Val Leu Ser Lys Asn Ile>

  580          590          600          610          620
     .    .       .    .       .    .       .    .       .
TTA AAA TCC GGA GAA ATA ACA GTT GCA CTT GAT GAC TCT GAC ACT ACT
AAT TTT AGG CCT CTT TAT TGT CAA CGT GAA CTA CTG AGA CTG TGA TGA
Leu Lys Ser Gly Glu Ile Thr Val Ala Leu Asp Asp Ser Asp Thr Thr>

    630          640          650          660          670
  .    .       .    .       .    .       .    .       .    .
CAG GCT ACT AAA AAA ACT GGA AAA TGG GAT TCA AAA ACT TCT ACT TTA
GTC CGA TGA TTT TTT TGA CCT TTT ACC CTA AGT TTT TGA AGA TGA AAT
Gln Ala Thr Lys Lys Thr Gly Lys Trp Asp Ser Lys Thr Ser Thr Leu>

       680          690          700          710          720
  .    .       .    .       .    .       .    .       .    .
ACA ATT AGT GTT AAC AGC AAA AAA ACT ACA CAA CTT GTG TTT ACT AAA
TGT TAA TCA CAA TTG TCG TTT TTT TGA TGT GTT GAA CAC AAA TGA TTT
Thr Ile Ser Val Asn Ser Lys Lys Thr Thr Gln Leu Val Phe Thr Lys>

          730          740          750          760
     .    .       .    .       .    .       .    .       .
CAA TAC ACA ATA ACT GTA AAA CAA TAC GAC TCC GCA GGT ACC AAT TTA
GTT ATG TGT TAT TGA CAT TTT GTT ATG CTG AGG CGT CCA TGG TTA AAT
Gln Tyr Thr Ile Thr Val Lys Gln Tyr Asp Ser Ala Gly Thr Asn Leu>
```

FIGURE 23 (2 of 3)

K48 OSPA / PGAU OSP A FUSION

```
770          780          790          800          810
 •      •     •      •     •      •     •      •     •
GAA GGC ACA GCA GTC GAA ATT AAA ACA CTT GAT GAA CTT AAA AAC G
CTT CCG TGT CGT CAG CTT TAA TTT TGT GAA CTA CTT GAA TTT TTG CG
Glu Gly Thr Ala Val Glu Ile Lys Thr Leu Asp Glu Leu Lys Asn Ala>

  820
    •     •
TTA AAA TAA
AAT TTT ATT
Leu Lys ***>
```


FIGURE 23 (3 of 3)

B-31 OSP A /PGAU OSP A FUSION

```
          10          20          30          40
     *    *    *    *    *    *    *    *    *
ATG AAA AAA TAT TTA TTG GGA ATA GGT CTA ATA TTA GCC TTA ATA GCA
TAC TTT TTT ATA AAT AAC CCT TAT CCA GAT TAT AAT CGG AAT TAT CGT
Met Lys Lys Tyr Leu Leu Gly Ile Gly Leu Ile Leu Ala Leu Ile Ala>

50          60          70          80          90
 *    *    *    *    *    *    *    *    *    *
TGC AAG CAA AAT GTT AGC AGC CTT GAT GAA AAA AAC AGC GCT TCA GTA
ACG TTC GTT TTA CAA TCG TCG GAA CTA CTT TTT TTG TCG CGA AGT CAT
Cys Lys Gln Asn Val Ser Ser Leu Asp Glu Lys Asn Ser Ala Ser Val>

  100         110         120         130         140
    *    *    *    *    *    *    *    *    *
GAT TTG CCT GGT GAG ATG AAA GTT CTT GTA AGT AAA GAA AAA GAC AAA
CTA AAC GGA CCA CTC TAC TTT CAA GAA CAT TCA TTT CTT TTT CTG TTT
Asp Leu Pro Gly Glu Met Lys Val Leu Val Ser Lys Glu Lys Asp Lys>

    150         160         170         180         190
 *    *    *    *    *    *    *    *    *    *
GAC GGT AAG TAC AGT CTA AAG GCA ACA GTA GAC AAG ATT GAG CTA AAA
CTG CCA TTC ATG TCA GAT TTC CGT TGT CAT CTG TTC TAA CTC GAT TTT
Asp Gly Lys Tyr Ser Leu Lys Ala Thr Val Asp Lys Ile Glu Leu Lys>

       200         210         220         230         240
    *    *    *    *    *    *    *    *    *
GGA ACT TCT GAT AAA GAC AAT GGT TCT GGA GTG CTT GAA GGT ACA AAA
CCT TGA AGA CTA TTT CTG TTA CCA AGA CCT CAC GAA CTT CCA TGT TTT
Gly Thr Ser Asp Lys Asp Asn Gly Ser Gly Val Leu Glu Gly Thr Lys>

          250         260         270         280
     *    *    *    *    *    *    *    *    *
GAT GAC AAA AGT AAA GCA AAA TTA ACA ATT GCT GAC GAT CTA AGT AAA
CTA CTG TTT TCA TTT CGT TTT AAT TGT TAA CGA CTG CTA GAT TCA TTT
Asp Asp Lys Ser Lys Ala Lys Leu Thr Ile Ala Asp Asp Leu Ser Lys>

290         300         310         320         330
 *    *    *    *    *    *    *    *    *    *
ACC ACA TTC GAA CTT TTA AAA GAA GAT GGC AAA ACA TTA GTG TCA AGA
TGG TGT AAG CTT GAA AAT TTT CTT CTA CCG TTT TGT AAT CAC AGT TCT
Thr Thr Phe Glu Leu Leu Lys Glu Asp Gly Lys Thr Leu Val Ser Arg>

  340         350         360         370         380
    *    *    *    *    *    *    *    *    *
AAA GTA AGT TCT AGA GAC AAA ACA TCA ACA GAT GAA ATG TTC AAT GAA
TTT CAT TCA AGA TCT CTG TTT TGT AGT TGT CTA CTT TAC AAG TTA CTT
Lys Val Ser Ser Arg Asp Lys Thr Ser Thr Asp Glu Met Phe Asn Glu>
```


FIGURE 24 (1 of 3)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/01424

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/11 C07K14/20 C12N15/62 C12N15/63 C12Q1/68
G01N33/569 A61K39/002

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FENG S ET AL: "CHARACTERIZATION OF TWO GENES, P11 AND P5, ON THE BORRELIA BURGDORFERI 49-KILO BASE LINEAR PLASMID" BIOCHIMICA ET BIOPHYSICA ACTA, vol. 1307, no. 3, 17 July 1996, pages 270-272, XP000613914	1-7, 16-24, 29,30, 33-36, 39, 43-46,48
Y	see the whole document	1-7, 16-24, 29,30, 33-36, 39, 43-46, 48, 50-56, 59-66

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 March 1999	26/03/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Ceder, O

1

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 95 35119 A (BOARD OF REGENTS, THE UNIVERSITY OF TEXAS) 28 December 1995	1-7, 16-24, 29,30, 33-36, 39, 43-46, 48, 50-56, 59,61, 62,65,66
A	see the whole document	25-28, 31,32
Y	EP 0 540 457 A (SYMBICOM AB) 5 May 1993 see abstract	60,63,64
A	WO 90 04411 A (SYMBICOM AB) 3 May 1990 cited in the application see abstract; claims 15,28-31,36	50,53
A	WO 95 12676 A (ASS UNIVERSITIES INC) 11 May 1995 see abstract; claims	33-35, 50,55
P,X	EMBL Data Base Empro:AE001117 Ac. No. AE001117, 16 December 1997 Fraser et al.: 'Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi' XP002096131 see the whole document	1-15
P,X	-& FRASER ET AL.: NATURE , vol. 390, 11 December 1997, pages 580-586, XP002096130	1-15, 20-32
P,X	EMBL Data Base Pir2:B70104 Ac. No. B70104, 13 February 1998 Fraser et al.: 'Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi' XP002096132 see the whole document	20-32
P,X	& FRASER ET AL.: NATURE, vol. 390, 11 December 1997, pages 580-586,	1-15, 20-32

1

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9535119 A	28-12-1995	AU 697305 B	01-10-1998
		AU 2945295 A	15-01-1996
		CA 2193682 A	28-12-1995
		EP 0766567 A	09-04-1997
EP 0540457 A	05-05-1993	AU 2903892 A	21-05-1993
		EP 0650527 A	03-05-1995
		WO 9308306 A	29-04-1993
WO 9004411 A	03-05-1990	AT 150318 T	15-04-1997
		AU 639667 B	05-08-1993
		AU 4495389 A	14-05-1990
		CA 2001328 A	24-04-1991
		DE 68927891 D	24-04-1997
		DE 68927891 T	02-10-1997
		DK 75091 A	12-06-1991
		EP 0445135 A	11-09-1991
		EP 0565208 A	13-10-1993
		EP 0711563 A	15-05-1996
		US 5523089 A	04-06-1996
		US 5777095 A	07-07-1998
		US 5582990 A	10-12-1996
		US 5688512 A	18-11-1997
WO 9512676 A	11-05-1995	AU 8127494 A	23-05-1995
		CA 2175567 A	11-05-1995
		EP 0726955 A	21-08-1996
		ZA 9408551 A	03-07-1995